

MICROCOPY RESOLUTION TEST CHART  
 NATIONAL BUREAU OF STANDARDS-1963-A

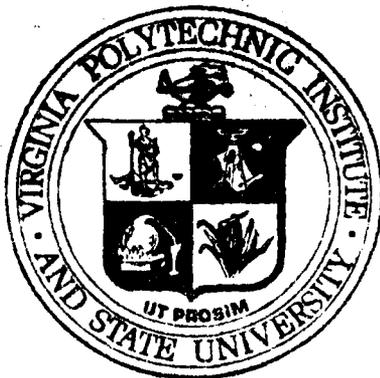
AD-A175 062

MINIARY SUBTHRESHOLD COLOR-DIFFERENCE  
METRICS OF LEGIBILITY  
FOR CRT PAPER IMAGERY

BY

Robert M. Lippert and Harry L. Snyder

OTIP FILE COPY



Virginia Polytechnic Institute  
and State University

Industrial Engineering and Operations Research  
BLACKSBURG, VIRGINIA 24061

86 12 11

UNITARY SUPRATHRESHOLD COLOR-DIFFERENCE  
METRICS OF LEGIBILITY  
FOR CRT RASTER IMAGERY

by

Thomas M. Lippert and Harry L. Snyder

Human Factors Laboratory  
Department of Industrial Engineering and Operations Research  
Virginia Polytechnic Institute and State University  
Blacksburg, Virginia 24061

September, 1986

DTIC  
ELECTE  
DEC 11 1986  
S D  
B

DISTRIBUTION STATEMENT A

Approved for public release;  
Distribution Unlimited

86 12 11 101

REPORT

# AD-A175 062

READ INSTRUCTIONS  
BEFORE COMPLETING FORM

1. REPORT NUMBER

RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

Unitary Suprathreshold Color-Difference Metrics of Legibility for CRT Raster Imagery

5. TYPE OF REPORT & PERIOD COVERED

Technical Report  
1982-1985

6. PERFORMING ORG. REPORT NUMBER

HFL/ONR 86-3

7. AUTHOR(s)

Thomas M. Lippert and Harry L. Snyder

8. CONTRACT OR GRANT NUMBER(s)

N00014-78-C-0238

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Human Factors Laboratory  
Virginia Polytechnic Institute and State University  
Blacksburg, VA 24061

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

NR 196-55

11. CONTROLLING OFFICE NAME AND ADDRESS

Office of Naval Research, Code 455  
800 N. Quincy Street  
Arlington, VA 22217

12. REPORT DATE

September, 1986

13. NUMBER OF PAGES

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)

Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Visual perception	Visual displays
Color vision	Legibility
Uniform chromaticity scales	CRT displays
	Visual detection

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This research examined the relationships between color contrast and legibility for digital raster video imagery. CIE colorimetric components were combined into three-dimensional color coordinate systems whose coordinates map one-to-one with the physical energy parameters of all colors. The distance between any two colors' coordinates in these 3-spaces is termed Color-Difference ( $\Delta E$ ).  $\Delta E$  was hypothesized as a metric of the speed (RS) with which observers possessing normal vision could accurately read random numeral strings of one color displayed

against backgrounds of another color.

Two studies totaling 32064 practice and experimental trials were conducted. The first study determined that the CIE Uniform Color Spaces are inappropriate for the modelling of RS. Subsequently, a different 3-space geometry and colorimetric component scaling were empirically derived from the Study 1 data to produce a one-dimensional  $\Delta E$  scale which approximates an interval scale of RS. This  $\Delta E$  scale and others were then applied to the different stimulus conditions in Study 2 to determine the generalizability of such  $\Delta E$  metrics.

The pair of studies is conclusive: several  $\Delta E$  scales exist which serve equally well to describe or prescribe RS with multicolor CRT raster imagery for a range of character luminances in both positive and negative presentation polarities. These are the  $Y, u', v'$ ,  $\log Y, u', v'$ ,  $L^*, u', v'$ , and  $L^*, u^*, v^*$  rescaled color spaces. Because of its predictive accuracy and simplicity, a luminance-generalized,  $\Delta E$ -standardized  $Y, u', v'$  metric, accounting for 71% and 75% of the RS variability in Studies 1 and 2, respectively, is recommended as the most appropriate metric of emissive display legibility to be tested in these studies.

## ACKNOWLEDGEMENTS

For the development and improvement of the Human Factors Laboratory's (HFL) display/radiometric measurement system and related software, great appreciation is expressed to Mr. Willard W. Farley, Dr. James C. Gutmann, Mr. Edward B. Costanza, and Dr. David L. Post.

The patience and support of Mr. Gerald S. Malecki, our technical monitor from the Office of Naval Research, are greatly appreciated, particularly for his understanding that this technical report is much more complete due to the inclusion of Study 2 results, obtained long after the completion of the subject contract.

Accession For	
NTIS GPO	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	<input type="checkbox"/>
By	
Distri	
Avail	
Dist	
A-1	

# TABLE OF CONTENTS

	page
ACKNOWLEDGEMENTS.....	i
TABLE OF CONTENTS.....	ii
LIST OF TABLES.....	iv
LIST OF FIGURES.....	v
INTRODUCTION.....	1
Overview.....	1
Color Science.....	4
Color additivity.....	5
Uniform Color Scales.....	13
Color-difference ( $\Delta E$ ).....	17
Previous color-contrast research.....	19
Development of a general color-difference metric of legibility.....	22
Chrominance axes rescaling.....	37
Research Objectives.....	46
METHOD.....	48
Digital Color Video System.....	48
Monitor.....	48
CRT convergence.....	49
Raster delay.....	50
Radiometric measurement system.....	52
Calibration.....	54

Sources of error.....	56
System stability.....	57
System baseline tests.....	57
Experiment control.....	57
Stimuli.....	58
Active display.....	58
Numeral strings and background.....	60
Interstimulus image.....	61
Experimental Design.....	62
Luminance and chrominance contrast.....	62
Experimental manipulations.....	62
Subjects.....	65
Procedure.....	65
RESULTS.....	70
APPLICATIONS.....	81
CONCLUSIONS.....	84
REFERENCES.....	88
<u>APPENDIX</u>	
A. Stimulus Color Parameters: Studies 1 and 2 (Lippert, 1984, 1985).....	92

LIST OF TABLES

TABLE	page
1. Common Names of Pure Spectral Hues.....	6
2. $\Delta E$ Component Weightings and Regression Coefficients for the Rescaled Metrics.....	71
3. $\Delta E$ Component Weightings and Luminance-Generalized, $\Delta E$ -Standardized Expressions of the Study 1 Metrics.....	77
4. Stimulus Color Parameters: Study 1.....	92
5. Stimulus Color Parameters: Study 2.....	95

## LIST OF FIGURES

FIGURE	page
1. Maxwell's color triangle.....	8
2. Grassman's tri-stimulus color space.....	9
3. CIE 1931 tristimulus space.....	11
4. The MacAdam (1942) ellipses.....	18
5. $\Delta E(L^*, u^*, v^*)$ by RS scattergram for $T_{ACH}$ showing least squares regression line .....	23
6. $\Delta E(L^*, u^*, v^*)$ by RS scattergram for $T_{Y-G}$ showing least squares regression line.....	24
7. $\Delta E(L^*, u^*, v^*)$ by RS scattergram for $T_{RED}$ showing least squares regression line.....	25
8. Cross-section of an arbitrary convergent color space in which chromatic planes are seen edgewise as horizontal line segments.....	27
9. $\Delta E(Y, u', v')$ by RS scattergram: $T_{RED}'$ $B_r, B_p$ removed Study 1 (Lippert, 1984).....	31
10. $\Delta E(Y, u', v')$ by RS scattergram: rescaled chrominance axes, Study 1 (Lippert, 1984).....	39
11. $\Delta E(\text{Log}Y, u', v')$ by RS scattergram: rescaled chrominance axes, Study 1.....	40
12. $\Delta E(L^*, u', v')$ by RS scattergram: rescaled chrominance axes, Study 1.....	41
13. $\Delta E(L^*, u^*, v^*)$ by RS scattergram: rescaled	

axes, (Study 1).....	43
14. Luminance-generalized, $\Delta E$ -standardized $\Delta E(Y,u',v')$ space by RS scattergram, Study 1.....	47
15. Black and white photographic examples of dynamic convergence of the shadowmask display.....	51
16. Digital color video system and measurement system block diagram.....	53
17. Black-and-white photographic example of HUD symbology, background fields, and the 3-, 4-, and 5-digit reading tasks.....	59
18. Chromaticities for HUD symbology and background fields on the CIE 1931 (x,y) chromaticity diagram...	61
19. Nominal experimental manipulations of numeral-to- background luminance modulation (M), background chromaticity (B), and number of reading task digits (N) for each of achromatic yellow-green, and red numeral chromaticities (T).....	63
20. Subject station.....	67
21. Luminance-generalized, $\Delta E$ -standardized $(Y,u',v')$ by RS scattergram, Study 2.....	74
22. Luminance-generalized, $\Delta E$ -standardized $(L^*,u^*,v^*)$ by RS scattergram, Study 2.....	75

## INTRODUCTION

### Overview

The recent proliferation of single- and multi-color electronic displays and display technologies is accompanied by a growing need to understand how best to employ them. Galves and Brun (1982) indicated the following potential advantages for multichrome over monochrome technologies in luminous data displays: increase in displayable information density, greatly reduced data acquisition time, greatly reduced risk of error in symbol and number identification, and the possibility of color-coding, supplementing information without requiring a shape-coding system.

Color producibility is restricted by the limitations of specific display technologies. Also, available colors are often displayed with primary concern for the subjective appreciation of display users, but not for objective, task-related visual performance measures. Although the improper or sub-optimal usage of display colors may be of little consequence in many applications, there exist military, industrial, and educational settings in which the considered usage of simultaneously displayed colors, generally in the form of alphanumeric and background fields, may be a deciding factor in the comfortable, timely, or even

successful completion of a mission or process the display is intended to support.

One critical example is the mission requirements of an aircraft equipped with a head-up display (HUD). Its electronic display-generated flight control or weapons delivery information is made available to the pilot in the form of visual symbols optically combined with the pilot's normal field of view through the windscreen. The design of a HUD is tested through many hours of in-flight operation. It is successful to the extent that a pilot can quickly and accurately extract (read) luminous information displayed against the real-world background--a dynamic, spatially complex visual field varying greatly in luminance and chrominance. For this and similar applications, a means to predict the performance effects of displayed color contrast would aid designers of multichrome display systems.

A study by Lippert (1984), which will be denoted "Study 1" throughout this report, focused on reading speed (RS) and accuracy in tasks requiring strings of dot-matrix numerals displayed against spatially uniform backgrounds to be read from a static, full-color simulated cathode-ray tube (CRT) HUD. RS is the reciprocal of the time required to accurately read a random numeral string (98% correct responses) and is therefore a proportional measure of improvement in, or ease of, legibility. The numeral reading

task is free of the language redundancy confounds associated with word reading tasks. Instead of the pilot's view through the windscreen, the HUD was programmed into a computer-generated real-world representation composed of 10 geometric fields corresponding to sky, clouds, and terrain, implying an horizon. Each of the independently addressable background fields could be assigned any color attainable with the selected color monitor. Therefore, the operational definition of color contrast for that study is the combination of luminance and chrominance difference between a given reading task numeral string, or target (T), and its immediate background field (B) which distinguishes one from the other both in terms of colorimetric parameters and RS. The investigation was undertaken to describe better the stimulus-response relationships of these two variables.

Study 2, described in detail in this report, is an investigation of the generalizability of displayed data legibility predictors, empirically derived from RS in Study 1, to different stimulus levels. The laboratory apparatus and experimental method are the same as for Study 1 and Study 2. In Study 2, reading speed as a function of color contrast was determined for achromatic, yellow-green, and red random numeral strings 3, 4, and 5 digits in length as in Study 1, but viewed at 0.76 m as opposed to 0.5 m in Study 1. The experiment also incorporated both positive

(e.g., lighter targets on a darker background) and negative (e.g., darker targets on a lighter background) polarities as opposed to Study 1's limitation to positive polarities only. Negative polarity was included to answer questions about color contrast formatting in head-down displays and other daylight displays not constrained to the HUD's positive polarity.

#### Color Science

Because the prediction of visual task performance from quantified color displays depends on color specifications, the following discussion is pertinent. The subject of color incorporates both the physical parameters of colors and the visual experience or sensations they produce. As a broad psychophysiological concept, color includes not only the sensation of hues, but also grays which, as characteristics of light, may be described objectively in terms of photometric quantity, dominant wavelength, and excitation purity. These three dimensions roughly correspond to the subjective visual attributes of brightness, hue, and saturation, respectively.

The most practical standard white light has traditionally been direct noon sunlight, although atmospheric pollution has necessitated its replacement with Standard Illuminants C

or D. White light is termed achromatic since it contains all visible wavelengths in proportions of equally perceived intensity.

Most colors are chromatic, however, in that they exhibit hue, which is directly related to the differences between their spectral energy distributions and that of sunlight (or an established substitute). The result of these differences is the sensation of a "greenish," "reddish," etc., hue. Color scientists have determined that, according to the average observer's judgment, the common names of pure spectral hues should be applied to the wavelength ranges in Table 1.

Color additivity. The Young-Helmholtz, or trichromacy, theory has been one relatively successful means of explaining many color vision phenomena. Developed at about the turn of the 19th century, it holds that the human eye is capable of three separate color sensations, corresponding to retinal stimulation by red, green, and blue wavelengths. These sensations combine so that every color sensation is the effect of the joint stimulation of three discrete retinal elements in some definite proportion. Red, green, and blue are thus known as the additive primary colors.

Many refinements have been made to this concept, which actually originated with Newton, who mechanically modeled

Table 1. Common Names of Pure Spectral Hues  
(Wyszecki and Stiles, 1967).

<u>COLOR</u>	<u>Wavelength, nm</u>
Violet	390-455 (complement)
Blue	455-492
Green	492-577
Yellow	577-597
Orange	597-622
Red	622-770

the primaries as weights placed at the vertices of a triangular plate. The resultant color perception was analogous to the center of gravity of the system, though Newton believed yellow was a primary color. Figures 1, 2, and 3 represent a progression of concepts to the present working model of color additivity.

Maxwell conceived the color triangle, clarifying the concept of color additivity with an equilateral triangle in which the altitude from each primary to the opposite side constitutes a value of 100. The proportionate lengths of the altitudes intersecting at any color within the triangle were the proportions of the respective primaries to be mixed to duplicate the color. For color H in Figure 1, the distances  $r$ ,  $g$ , and  $b$  represent the proportions of the primaries R, G, and B which, when added together, will imitate H. Removing a primary color's spectral energy from full-spectrum white light yields the subtractive complement shown opposite it (i.e., yellow results from the subtraction of blue from white).

Grassman (1853) formalized the laws of color mixture in three-dimensional tristimulus space by intersecting the color triangle at points of equal perceptual intensity along the primaries' intensity vectors, R, G, and B, creating the unit (intensity) plane, or chromaticity diagram (Figure 2). Tristimulus space is useful because most colors may be

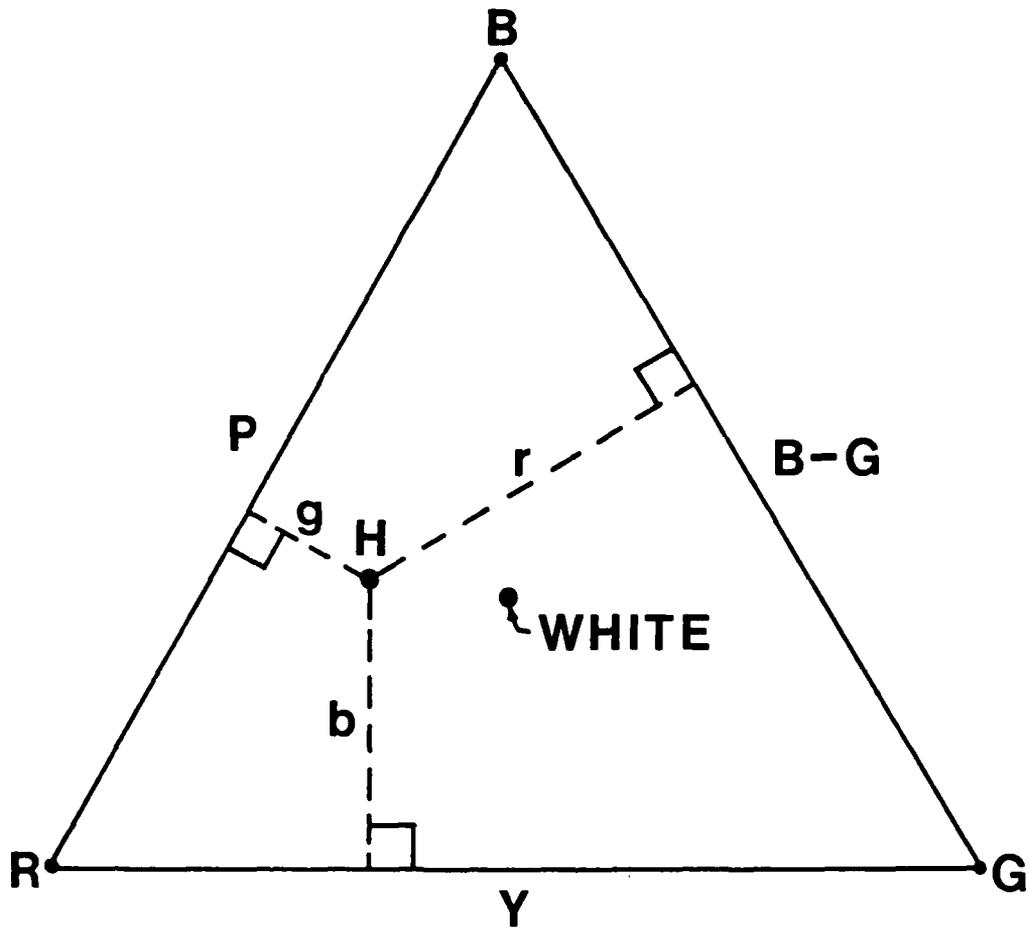


Figure 1. Maxwell's color triangle.

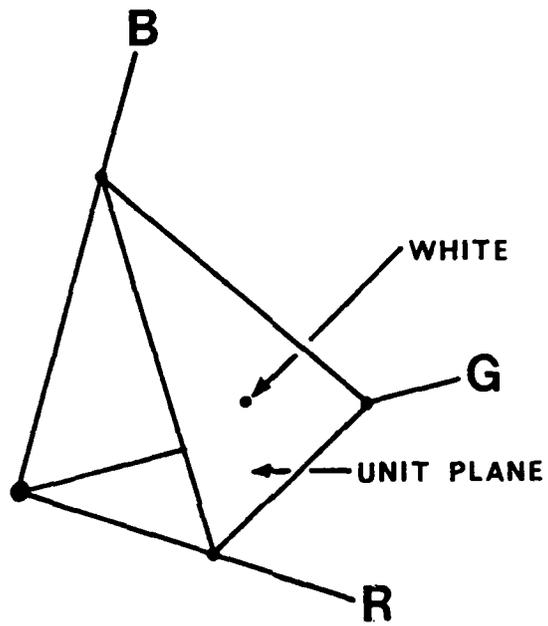


Figure 2. Grassman's tristimulus color space.

specified in terms of the simple additive combination of quantities of each primary constituting them (e.g., the tristimulus values of a color of interest). The added tristimulus values of two colors specify a third color which is a perceptual match to the mixture of the first two. This innovation made the quality control of mixing, matching, and maintaining the colors of pigmented or reflective sources more precise, easily communicated, and even routine.

However, color-matching experiments (see Graham, 1965, p. 370), using the red, green, and blue primaries (at wavelengths 700, 546.1, and 435.8 nm: the "RGB System"), consistently showed that one of the reference (primary) stimuli had to be added to the test (to-be-matched) stimulus, effectively desaturating it, in order for a match to be made of some test colors. In other words, there were some (pure) colors which could not be matched through the additive mixture of the primaries, and this left the science of color additivity incomplete.

A solution was found by the Commission Internationale de l'Eclairage (CIE) in 1931 by specifying imaginary primaries such that all real colors could be achieved through their additive combination. This system uses the labels X, Y, and Z for the primary vectors, creating the chromaticity diagram seen in Figure 3, which describes normal human trichromatic additivity relative to these mathematical primaries. The



Standard Observer for Colorimetry (1931 2-degree Standard Observer) is defined by this convention, derived from data from two color matching studies (Guild, 1931; Wright, 1928-29), which are considered the best-controlled to date and which provide the current colorimetric standard.

All pure spectral colors lie on the spectrum locus, except for the purples across the bottom of the space, which are all additive combinations of blues and reds and which may be specified in complementary wavelengths--the negative of the spectrum locus wavelength colinear with the purple stimulus chromaticity coordinates and E, or Standard Illuminant C, the neutral reference color. The ratios of each of the tristimulus values, X, Y, Z, to their sum,  $(X + Y + Z)$ , are the chromaticity coordinates, x, y, z, which thus add to unity for every color. They specify a color's dominant wavelength and excitation purity (e.g., the proportionate colinear distance from E to the spectrum locus for a color), but do not indicate its perceived intensity, or brightness.

Earlier work by Gibson and Tyndall (1923) and Coblentz and Emerson (1917) had determined the physical intensities of many pure spectral hues required to match the brightness of a white stimulus of known physical intensity. This resulted in the establishment of the relative luminosity function of wavelength (e.g., CIE 1924 photopic luminosity

function). The spectral energy distribution of any light source may be weighted by the luminosity function, then integrated across its spectrum, yielding the measure termed luminance (L). Tristimulus space was constructed such that Y may be scaled to represent luminance (more will be said of the empirical relationship between brightness and luminance later). Y, x, and y are defined to be mathematically orthogonal and may therefore be combined to form a three-dimensional coordinate space comprised of a chromaticity diagram normal to a luminance axis.

Uniform color scales. While tristimulus space and its Y,x,y variant serve as satisfactory reference constructs of the physical parameters of colors, their axes are not perceptually independent. Moreover, they were not designed to be perceptually uniform. Perceptual uniformity would be achieved by a space in which a unit of linear distance, regardless of its location or orientation with respect to the axes, represented an invariant measure of effective color contrast or perceptual difference. (The uniform space sought in the present research is one in which distance represents an interval scale of perceived color contrast measured indirectly by RS.)

Since the establishment of tristimulus space, much work has gone into the adjustment of the geometrical relationships among its color coordinates in order to render

it perceptually uniform with regard to a restricted set of color vision phenomena (i.e., color-matching behavior). In 1976, the CIE adopted the UCS (Uniform Color Scale) as an improvement over the x,y chromaticity diagram in mapping color saturation data. The UCS Psychometric Chromaticity Coordinates,  $u'$  and  $v'$ , are derived from CIE 1931 tristimulus values by the equations,

$$u' = 4X / (X + 15Y + 3Z), \text{ and} \quad (1)$$

$$v' = 9Y / (X + 15Y + 3Z). \quad (2)$$

Also in 1976,  $L^*$ , or CIE metric lightness, was accepted as a uniform perceptual scale of the intensity of colors derived from the Munsell Color System "Value" dimension, and is defined by the equation

$$L^* = 116(Y / Y_n)^{1/3} - 16, \quad (3)$$

$$(Y / Y_n) > 0.01,$$

where  $Y$  is the Y-tristimulus value of a given color and  $Y_n$  is that of a reference white. Because  $L^*$  was developed for surface colors (reflective surfaces), the value of  $Y_n$  was generally set at the reflectance of white, which is maximally 100%. (For the present studies,  $Y_n$  is arbitrarily set at 100 for the chromaticity coordinates of Standard Illuminant C.)

An early developed color space is CIE 1976 ( $L^*$ ,  $u^*$ ,  $v^*$ ), in which

$$u^* = 13L^* (u' - u'_n), \text{ and} \quad (4)$$

$$v^* = 13L^* (v' - v'_n), \quad (5)$$

where  $L^*$  is metric lightness,  $u'$  and  $v'$  are the UCS coordinates of a color, and  $u'_n$  and  $v'_n$  are the coordinates of the reference white.  $L^*$  is seen here to effect a linear projection of the chromatic space as a function of its own value, while the 13 corresponds to the 13 lightness levels in the Munsell value scale.

A second color space is CIE 1976 ( $L^*$ ,  $a^*$ ,  $b^*$ ), in which

$$a^* = 500((X / X_n)^{1/3} - (Y / Y_n)^{1/3}), \text{ and} \quad (6)$$

$$b^* = 200((Y / Y_n)^{1/3} - (Z / Z_n)^{1/3}), \quad (7)$$

$$(X / X_n), (Y / Y_n), (Z / Z_n) > 0.01,$$

where  $X$ ,  $Y$ , and  $Z$  are the tristimulus values of a color and  $X_n$ ,  $Y_n$ , and  $Z_n$  are those of the reference white.

As seen above, metric lightness is the achromatic intensity scale employed with either ( $u^*$ ,  $v^*$ ) or ( $a^*$ ,  $b^*$ ) chromaticity in the CIE 1976 uniform color spaces. These two variants of metric chromaticity differ in that ( $u^*$ ,  $v^*$ ) is a linear projection of 1931 ( $X, Y, Z$ ) space and straight

lines in  $(x,y)$  correspond to straight lines in  $(u^*,v^*)$  when  $L^*$  is held constant, while  $(a^*,b^*)$  is a curvilinear projection of  $(X,Y,Z)$  and straight lines in  $(x,y)$  are generally curved in  $(a^*,b^*)$ . In other words, points of constant dominant wavelength, the physical correlate to hue, fall along straight lines in  $(x,y)$  or  $(u^*,v^*)$ , but along curved lines in  $(a^*,b^*)$ . The existence of two "uniform" chromatic space designs stems from disparate findings in basic work on uniform color-spacing structures, or line element geodesics (Wyszecki and Stiles, 1982). The  $(u^*,v^*)$  spacings are recommended for large chromatic differences and the  $(a^*,b^*)$  spacings for small chromatic differences, yet a presumption of either form of metric chromaticity as more appropriate for the present research application is empirically unfounded.

A third proposed color space is  $(W, a, b)$  (Cohen and Friden, 1975; 1976). Unlike the  $L^*,u^*,v^*$  and  $L^*,a^*,b^*$  spaces, both transforms of the CIE 1931 tristimulus space,  $W,a,b$  coordinates may be calculated directly from any valid color-matching functions (such as tristimulus values) by a linear algebraic matrix manipulation procedure. (See Costanza, 1981, for complete computational details. W. W. Farley of Virginia Tech has derived the transform matrix for the CIE 1931 2-degree color-matching functions  $(x, y, z)$  required to specify experiment colors in  $W, a, b$  space.)

### Color-difference ( $\Delta E$ )

Color contrast may be considered an additive composite of luminance contrast and chrominance contrast. Luminance contrast (e.g., black-and-white, or achromatic) is specified here in terms of luminance modulation (M),

$$M = (L_{\max} - L_{\min}) / L_{\max} + L_{\min} \quad (8)$$

for a given target/background (T/B) combination, where  $L_{\max} = L_T$  and  $L_{\min} = L_B$ . Luminance contrast may be combined with chrominance contrast to create a measure of overall color contrast termed color-difference (e.g.,  $\Delta E$ ), which is computed using a Pythagorean distance formula,

$$\Delta E(L^*, u^*, v^*) = ((\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2)^{0.5}, \quad (9)$$

where  $\Delta u^*$  and  $\Delta v^*$  are the differences (distances) between the stimulus pair's transposed chromaticity coordinates and  $\Delta L^*$  is the difference between transposed luminances in the CIELUV space. Similar distance formulas can be calculated for other color spaces.

It is desirable to control stimulus chromaticities to within one just noticeable difference (JND) of specification over the duration of any experiment. In this regard, the work of MacAdam (1942), in which the CIE 1931 chromaticity diagram was explored to determine "equal noticeability

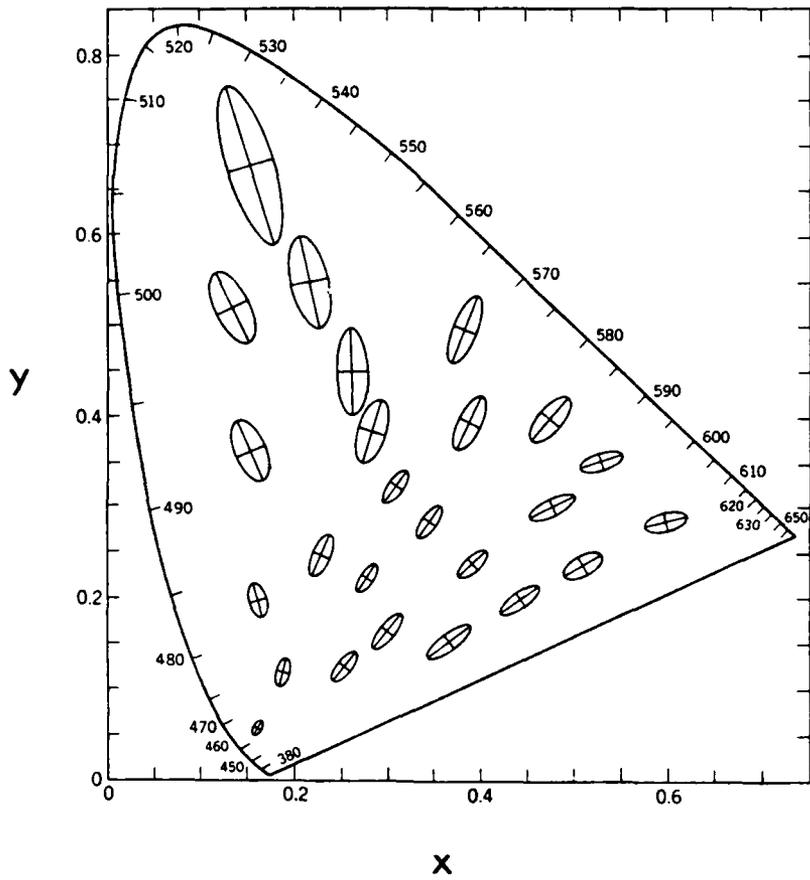


Figure 4. The MacAdam (1942) ellipses.

ellipses" for 25 chromaticities, shown in Figure 4, has been consulted. Colors at given ellipse boundaries were just noticeably different (based on the statistical variability of observer responses) from the fixed chromaticity at the intersection of the ellipses' axes for a single observer with normal color vision viewing a 2-degree test spot at 50 cd/m<sup>2</sup>. The ellipses are plotted in (x,y) coordinates times 10 actual size. The JND values ranged from 0.00058 to 0.00697 (x or y). These are the best available data from which to specify stability tolerances for stimulus chromaticities for the present research, yet the attainability of such tolerances with the present experimental apparatus is not readily testable, as discussed in the Method section.

Previous color contrast research. The difference between the brightness (e.g., phenomenal intensity) of a target and that of its background will determine to some extent their distinctiveness, or perceived contrast. As mentioned earlier, however, existing color spaces only approximate the expected brightness effect of colors through their scaling of luminance, or, in the case of W,a,b, achromatic intensity, providing poor-to-fair brightness metrics. The perceived brightness of a color is due not only to its luminance, but also to its chrominance. Studies conducted by Booker (1981) and others have shown that, for stimuli of

equal luminance, as the purity of a color increases, so does its perceived brightness. The magnitude of this perceptual phenomenon also varies with the dominant wavelength of a stimulus, being smallest for yellows and greatest for blues.

Our first goal in a long-term color contrast research program was to conceive an experimental procedure free of the wavelength-related brightness phenomenon. Subjects brightness-matched seven CRT chromaticities of dominant wavelengths from around the visible spectrum to 35, 50, and 70  $\text{cd/m}^2$  achromatic stimuli, resulting in three sets of equally bright colors used in subsequent experiments (see Costanza, 1981). The pairing of any two brightness-matched hues therefore resulted in a perceived contrast attributable to chromatic differences alone. These chromatic differences were scaled in terms of equivalent achromatic contrast by having observers adjust the luminances of two achromatic stimuli until their perceived contrast equalled that of each possible pair of brightness-matched colors. A second study paired all the color stimuli rather than only brightness-matched ones. It was reasoned that a color contrast metric of high utility must account for the hue-related brightness phenomenon, and that mean luminance modulation settings would correlate with color pair vector differences in such a uniform color space.

Evaluations of the  $L^*,u^*,v^*$ ,  $L^*,a^*,b^*$ , and  $W,a,b$  color spaces showed that 36% to 90% of the variation in mean modulation settings was accounted for by linear regression models of color contrast, with no evidence of higher-order effects (Post, Costanza, and Lippert, 1982). The predictive power of each metric was reduced in the study pairing all colors. A modified version of  $L^*,u^*,v^*$  yielded a superior predictor in all cases. The findings suggest that a useful transformation of color contrast into equivalent perceived achromatic contrast might be determined through further study, and that the equivalent achromatic contrast might be employed to specify the appropriate display of colors based on existing knowledge of visual performance with achromatic displays. The next research objective was to study the effects of color contrast in a more applied setting.

Lippert (1984) conducted the following visual performance study (i. e., Study 1). Ten normal trichromats' numeral string reading speed (RS) was determined for achromatic, yellow-green, and red dot-matrix numerals (T) presented against eight spatially uniform background chromaticities (B) on a 0.5 m distant high-resolution digital shadowmask display under low ambient illumination. The numeral colors were displayed at a constant  $47 \text{ cd/m}^2$  and paired with each background chromaticity displayed at seven luminance levels from 23 to  $47 \text{ cd/m}^2$ , resulting in a factorial design of

eight levels of chrominance contrast by seven levels of luminance modulation (M), or 56 color contrast conditions, for each numeral color. The complete listing of Study 1 stimulus color parameters is in Table 4, Appendix A. All color contrast conditions were tested for 3, 4, and 5 digit random numeral strings.

Performance with the achromatic and yellow-green numerals was similar overall, varying as a function of chrominance contrast but being dominated by a positive correlation with luminance contrast. RS asymptoted within the luminance contrast range tested. The red and purple backgrounds proved exceptions in that relatively high performance was obtained even without luminance contrast. The red numeral chromaticity overwhelmed both chrominance and luminance contrast effects, resulting in superior performance overall and further demonstrating the potential for chromatic contrast alone as a prime determinant of visual performance. The effect of the number of digits read, N, was shown to be independent of color contrast and linear over the range tested.

Development of a general color-difference metric of legibility. As described above, the first  $\Delta E$  metrics of RS examined were  $L^*, u^*, v^*$ ,  $L^*, a^*, b^*$ , and Cohen and Friden's 1976 W,a,b. As shown in Figures 5, 6, and 7, which are scattergrams of  $\Delta E$  versus mean RS ( $n = 90$ ) for  $T_{ACH}$ ,  $T_{Y-G}$ ,

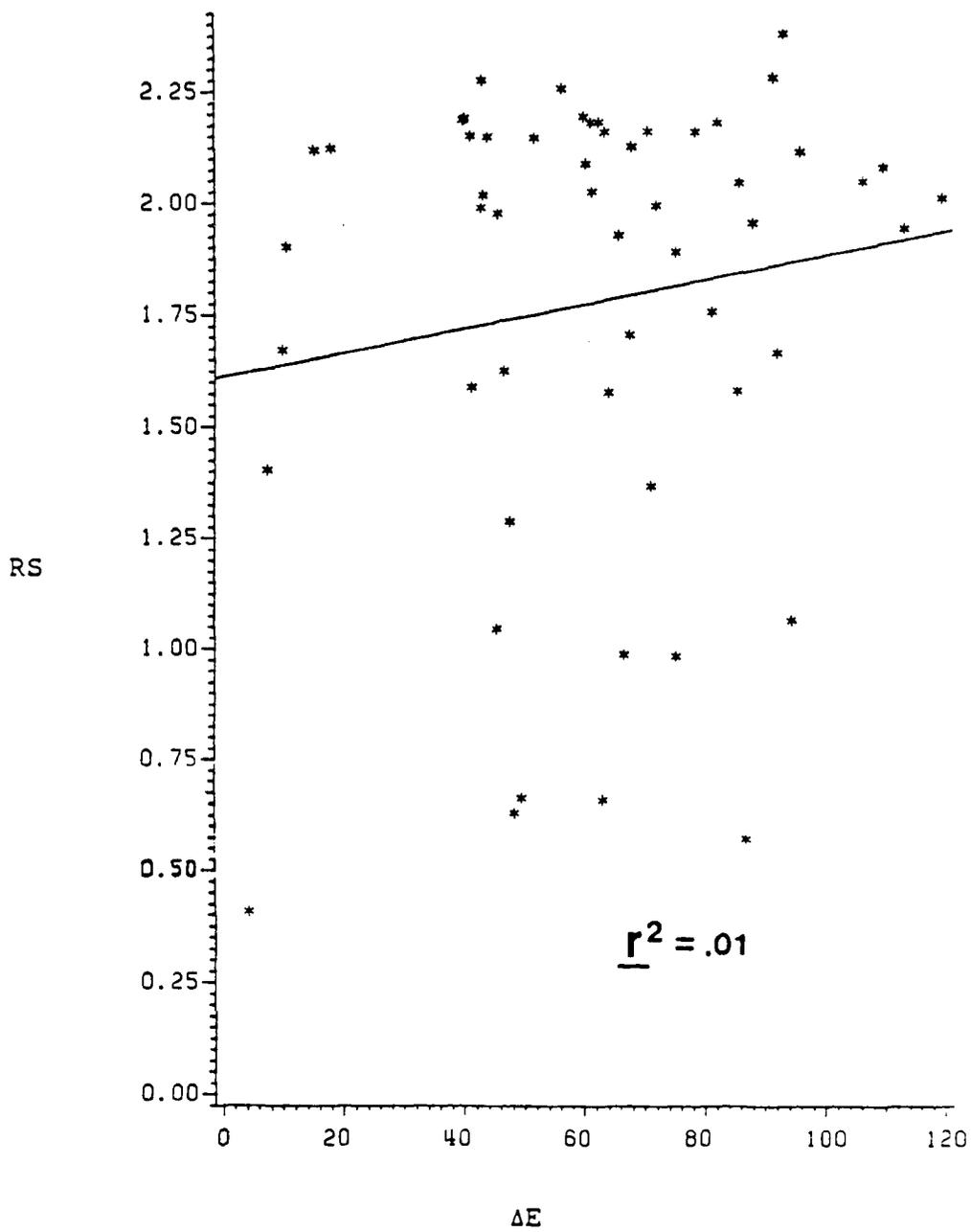


Figure 5.  $\Delta E$  by RS scattergram for  $T_{ACH}$  in  $L^+, u^+, v^+$  showing least squares regression line.

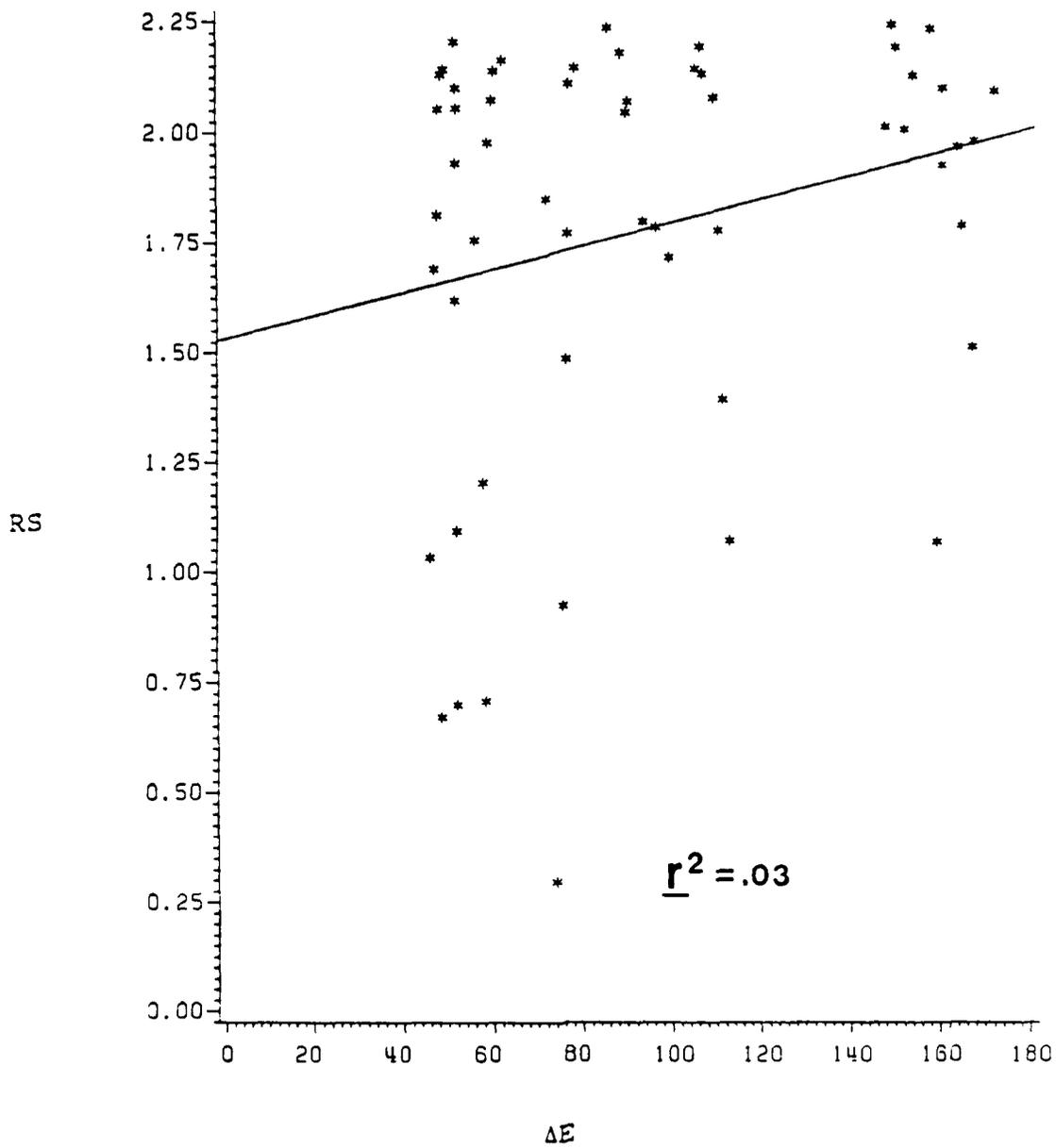


Figure 6.  $\Delta E$  by RS scattergram for  $T_{Y-G}$  in  $L^*, u^*, v^*$  showing least squares regression line.

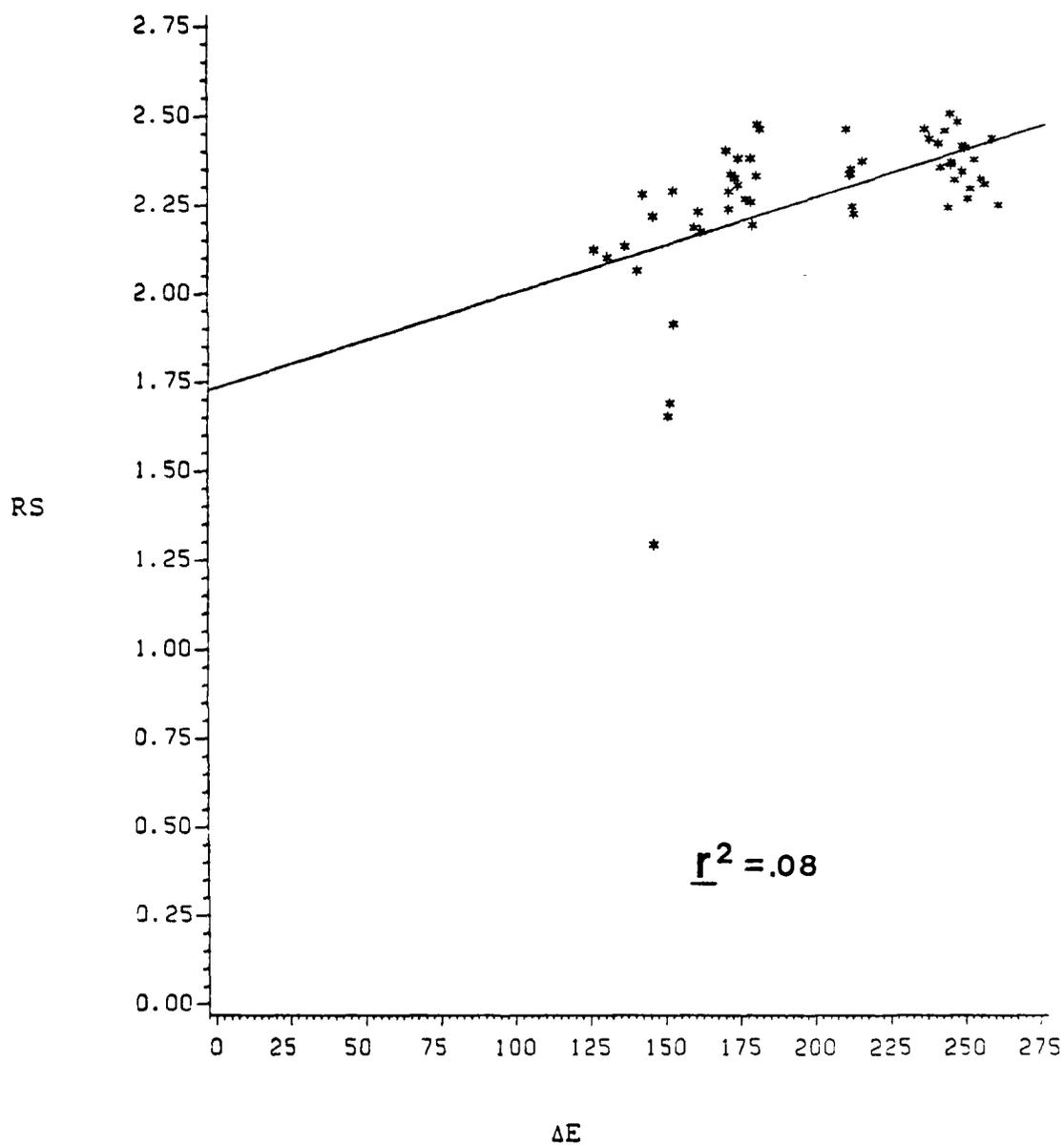


Figure 7.  $\Delta E$  by RS scattergram for  $T_{RED}$  in  $L^*, u^*, v^*$ .  
 showing least squares regression line.

and  $T_{RED}$ , respectively, with all of the  $B_s$  tested in Study 1,  $\Delta E$  in  $L^*, u^*, v^*$  is essentially unrelated to RS. A metric of legibility based on  $\Delta E$  in  $L^*, u^*, v^*$  is contra-indicated.

The primary source of the poor correlations between  $\Delta E$  and RS is the "convergent" nature of each of the color spaces investigated, shown for an arbitrary color space, in which any chromatic plane is seen edgewise as a horizontal line segment, and achromatic intensity is represented by the vertical dimension, as in Figure 8. Points 1 and 2 are arbitrary chromaticities, representing a target (T) and a background (B) of equal luminance for purposes of illustration, and the entire color gamut converges to the black point origin, O. The vertical axis M corresponds to the magnitude of luminance modulation. In the case of  $T_2 \times B_1$ , as M increases,  $B_1$ 's three-dimensional position in the space will proceed down  $L_1$  toward O, while the corresponding  $\Delta E$  decreases from point 1 to point 3, then increases beyond point 3 to O. Because RS increases monotonically with M, a color space resulting in such a non-monotonic computation of  $\Delta E$  must create problems, not only in fitting RS, but also in the application of any metric to specify color contrast, given a desired performance level and fixed T/B chromaticities, because there will be an M-range for many T/B combinations over which a distance specification (e.g., required  $\Delta E$  computed from the metric) will have two

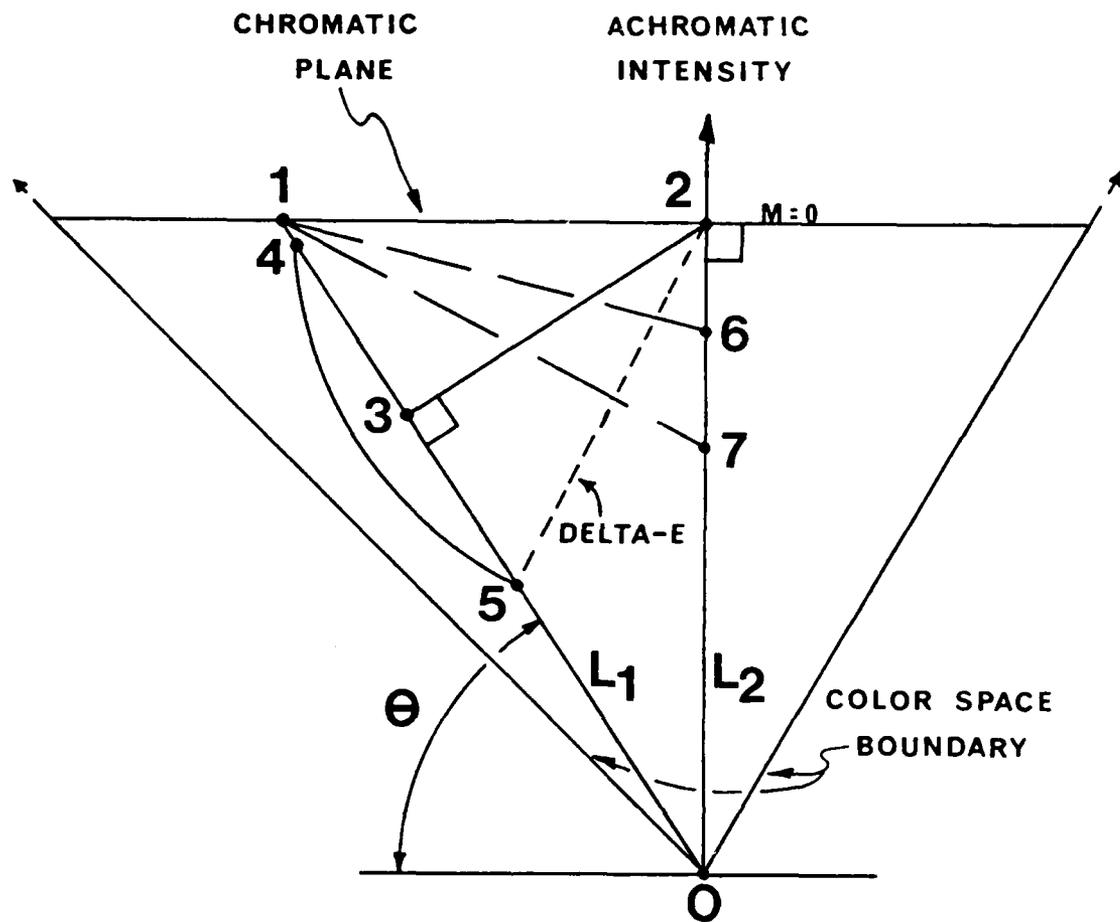


Figure 8. Cross-section of an arbitrary convergent color space in which chromatic planes are seen edgewise as horizontal line segments.

equidistant loci in the space along the increasing M dimension. This insurmountable property is illustrated in Figure 8 by points 4 and 5, equidistant from point 2. There is no logical approach to choosing between 4 and 5 working from the metric; thus, any such space is seen as invalid for the purpose of color production specification.

Actual  $\Delta E$  computations of Study 1 contrast combinations (Table 4, Appendix A) further illustrate the RS modeling inconsistencies of  $L^*, u^*, v^*$  space. Consider the  $T_{ACH} \times B_p$  combination versus the  $T_{RED} \times B_p$  combination over the range of M tested: at  $M_{.006}$ , computed  $\Delta E$  for  $T_{ACH} \times B_p$  is 119.6, at  $M_{.060}$  it is 113.3, at  $M_{.120}$  it is 109.7, then 106.3, 95.8, 93.0, and finally 91.4 at  $M_{.316}$ , describing an inverse relationship where distance in  $L^*, u^*, v^*$  decreases as M (with increasing RS) increases. This is a striking example of poor RS modeling when one realizes that at  $M_{.316}$ , with distance still decreasing, the point in the example which is analogous to the minimum distance point (3) in Figure 8 has not yet been reached for this T/B pair. The implication is that minimum  $\Delta E$  for  $T_{ACH} \times B_p$  (and therefore minimum RS) occurs at an M level higher than any tested in Study 1.

The reverse situation is seen to occur for  $T_{RED} \times B_p$  over M in  $L^*, u^*, v^*$ :  $\Delta E$  at  $M_{.006}$  is 146.3, at  $M_{.060}$  it is 151.0, at  $M_{.120}$  it is 151.5, then 152.9, 160.1, 161.4, and 162.1 at  $M_{.316}$ . This direct relationship between  $\Delta E$  and M is a more

reasonable descriptor of RS, although the distance in this case will continue to increase beyond the M level of maximum performance, suggesting that some non-linear relationship between RS and  $\Delta E$  will prove the best performance descriptor/predictor.

Obviously, there is an infinite number of T/B combinations specifiable within any color space and, as Figure 8 is viewed, it becomes clear that the trans-luminous range over which  $\Delta E$  decreases from the  $M = 0$  level is a complex function of the slope of the line of constant chromaticity for any B and the relative positions of both T and B with respect to point 2, or  $L_2$ , which is normal to the chromaticity plane and defines the neutral referent chromaticity, which, in the case of this study, is Standard Illuminant C.

The arbitrary model described above is most similar to  $L^*, u^*, v^*$  space. While  $L^*, a^*, b^*$  and  $W, a, b$  are somewhat different, their geometrical cases need not be specifically addressed here. All are convergent color spaces and yield similar inconsistencies in computed  $\Delta E$  with respect to RS.

A monotonic description of RS is obtained with a variation of the  $\Delta E$  formulation,

$$\Delta E_{40}(Y, u', v') = ((\Delta Y)^2 + (40 \Delta u')^2 + (40 \Delta v')^2)^{0.5}. \quad (10)$$

Because  $Y$ ,  $u'$ , and  $v'$  are physically as well as mathematically orthogonal, the Pythagorean  $\Delta E$  formula (10) defines a plane-origin, or "non-convergent", three-space in which the CIE luminance and chrominance constructs are orthogonally represented in all cases, unlike  $L^*, u^*, v^*$  for instance, in which  $u^*$  and  $v^*$  are linear projections of  $u'$  and  $v'$ , respectively, as a function of  $L^*$ .

Various luminance and chrominance units were tested in (10), indicating that CIE 1931  $Y$  ( $\text{cd}/\text{m}^2$ ) and CIE 1976 ( $u', v'$ ) UCS coordinates provided the best  $\Delta E$  descriptor of RS (Lippert, 1984). A weighting factor, 40, shown in (10), was statistically determined to adjust the relative contributions of luminance contrast and chrominance contrast component scales, which have no a priori relationship to one another, yielding a maximum linear  $r^2 = .64$  for the Study 1 data averaged across  $N$  and excluding red and magenta chromaticities. T/B combinations including red or magenta produced near-maximum RS without luminance contrast and are poorly described by this  $Y, u', v'$  metric. However, red and/or magenta combinations yield RS (e. g., legibility) which is underpredicted by the metric. Therefore, the metric is expected to be a very conservative predictor of RS for combinations of, or with, red or magenta.

Figure 9 is the  $\Delta E$  by mean RS scattergram for 83 color combinations excluding red and magenta ( $n = 90$ ) in

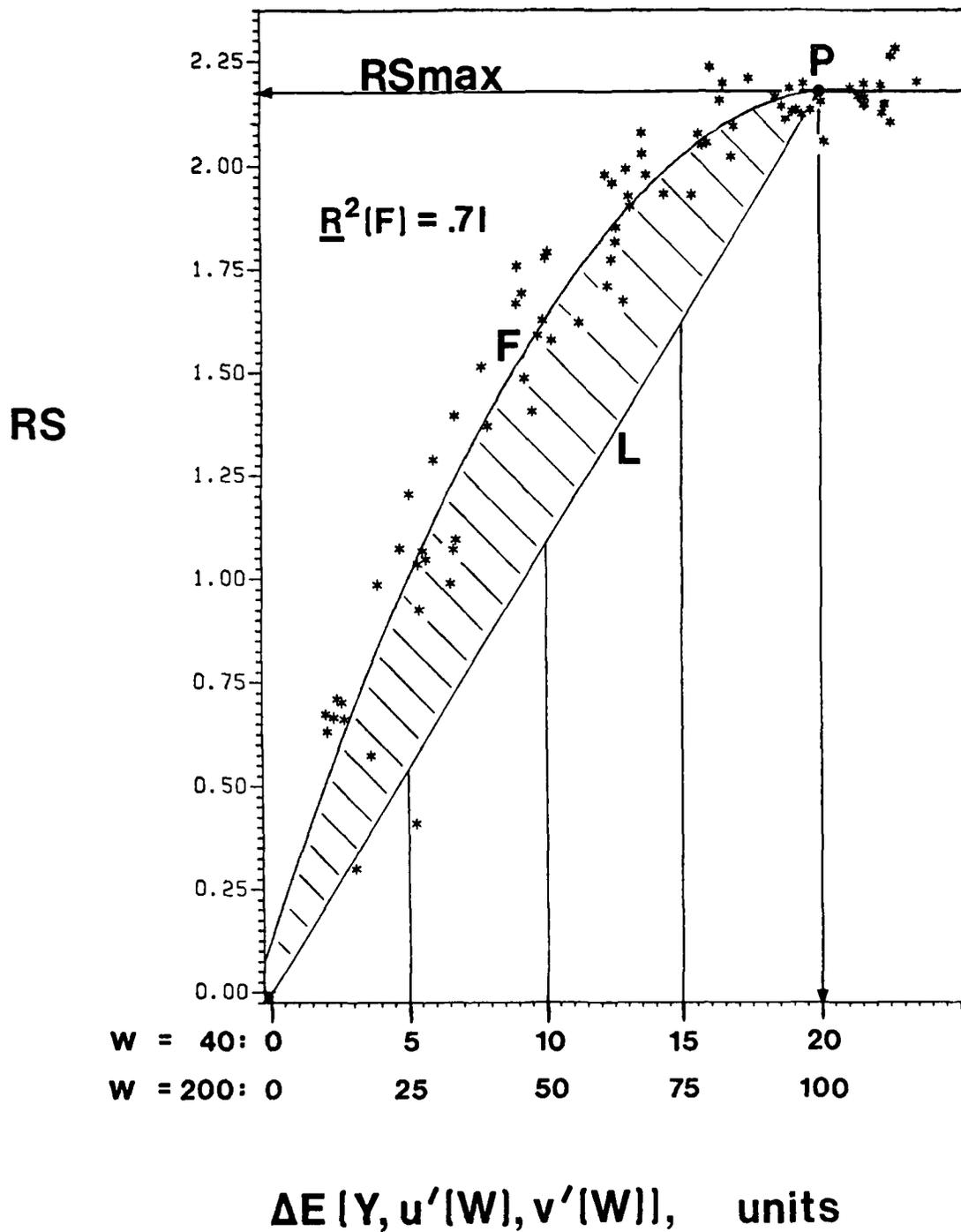


Figure 9.  $\Delta E(Y, u', v')$  by RS scattergram:  $T_{RED}$ ,  $B_r$ , and  $B_p$  removed, Study 1 (Lippert, 1984).

$Y, u'(40), v'(40)$  space. A parabola comprised of  $\Delta E$  and  $\Delta E^2$  terms has been fitted to the data ( $R^2 = .71$ ), revealing a performance asymptote ( $RS_{max} \approx 20 \Delta E$  units contrast) and a possible breakdown of the metric at low  $\Delta E$  based on the contrast conditions sampled. Such a failure of the metric is probably not of practical concern because such low  $\Delta E$  levels of contrast should never be prescribed for data display, although pictorial imaging may require the display of luminance and chrominance gradients of the smallest steps attainable with a display system in order to achieve more natural appearing representations of real-world objects.

If possible, we wish to generalize the empirical modeling results to the entire luminance range attainable with color raster CRTs. To begin this process, a simplifying assumption has been made. RS is assumed to be determined by color difference independent of absolute levels of photopic retinal stimulation. For example, T/B contrast comprised of a 1.5:1 luminance ratio and a 0.2 ( $u', v'$ ) unit difference is assumed to yield the same RS for any  $Y_T$  if the presentation polarity is positive or for any  $Y_B$  if the polarity is negative. The luminance ratio =  $Y_{max}/Y_{min}$ . Therefore, the scaling procedure is accomplished in all cases using  $Y_{max}$  ( $Y_M$ ), where  $Y_M$  is the mean target luminance. It will be shown immediately following description of an initial general metric how the metric might be adjusted to account

for variation in the performance contributions of luminance contrast relative to chrominance contrast as a function of  $Y_M$ .

The weighting factor, 40, is specific to the particular  $Y_M$  in Study 1, for which averaged  $Y_M = 45.9 \pm 0.5 \text{ cd/m}^2$ . When applying this  $\Delta E$  metric to the more broad  $Y_M$  continuum, the weighting factor must be scaled proportionately to the difference between  $Y_M = 45.9 \text{ cd/m}^2$  and the  $Y_M$  of interest in order to maintain the linear relationship between luminance ratio and unit chrominance difference, as RS determinant components of color contrast, characterized by  $\Delta E$  in (10).

Scaling of the weighting factor may be performed on a "per color pair" basis for any  $Y_M$  using the ratio of the (linearly related) optimal weighting factor and averaged  $Y_M$ s from Study 1,

$$(\text{weighting factor} / Y_M) = (40 / 45.9) = 0.87 = R, \quad (11)$$

where  $R$  is the assumed constant relationship between luminance contrast and chrominance contrast components. This results in a  $Y_M$ -corrected weighting factor,  $W$ , where,

$$0.87Y_M = W. \quad (12)$$

The meaning of 1 unit of  $\Delta E$  with respect to RS is specific to the weighting factor, 40, and the experimental  $Y_M$ s. However, if, as assumed, R is constant, it is possible to select any desired scaling of the  $\Delta E$  unit to provide the most convenient color difference scale for a given application. This is merely a standardization procedure.

As stated above for the Study 1 data in  $Y, u'(40), v'(40)$  space,  $\Delta E_{RSmax_{40}} = 20$ . To standardize the metric,  $\Delta E_{RSmax}$  was set to equal 100 units instead of 20. Then, a T/B color pair represented by 50  $\Delta E$  units is predicted to yield 50% of RSmax working from the standardized space. This feature is incorporated by scaling each of the three  $\Delta E$  components ( $\Delta Y$ ,  $\Delta u'$ , and  $\Delta v'$ ) with an appropriate ratio, S, determined by the fixed relationship between Study 1's weighting factor, 40, and RSmax equivalent  $\Delta E$ , 20,

$$40/20 = 2. \quad (13)$$

In the resulting legibility performance space, the standardized weighting factor becomes,

$$(\Delta E_{RSmax_{std}})(40 / 20) = 100 (2) = 200 = W_{std}, \quad \text{and} \quad (14)$$

$$W_{\text{std}} / W = S. \quad (15)$$

Instead of scaling  $\Delta E$  up or down based on  $Y_M$ , it is simpler, starting with  $\Delta Y$ ,  $\Delta u'$ , and  $\Delta v'$ , to scale any  $Y_M$  directly to the standardized  $Y, u', v'$  metric using the equations already described but shown now in correct order:

$$W = 0.87 (Y_M), \quad (16)$$

$$S = 200 / W, \quad \text{and} \quad (17)$$

$$\Delta E_{\text{std}}(Y, u', v') = ((S\Delta Y)^2 + (SW\Delta u')^2 + (SW\Delta v')^2)^{0.5}, \quad (18)$$

or simply,

$$(((230/Y_M)\Delta Y)^2 + (200\Delta u')^2 + (200\Delta v')^2)^{0.5}. \quad (19)$$

Equation (19) is all that is required to establish a standardized color difference metric of legibility for all  $Y_M$  in which  $\Delta E_{\text{RSmax}} = 100$ . Regression analyses performed using this  $\Delta E$  scale on different data sets with the same dependent measure will result in directly comparable b-coefficients. Of course, b-coefficients from  $Y, u', v'_{\text{std}}$  must be multiplied by  $S$  to return them to their original effect-per-unit meaning.

The Study 1 data were transformed to  $Y, u', v'_{\text{std}}$  with identical results to the original 40-weighting regression

procedure, confirming the basic rescaling relationships among  $R$ ,  $W$ ,  $W_{std}$ , and  $S$ .

Returning to Figure 9, if the  $RS_{max}$  equivalent  $\Delta E$ , 20, is rescaled to 100, and a straight line,  $L$ , is drawn from  $(0,0)$  to the  $RS$  asymptotic point,  $P$ , then the  $\Delta E$  scale becomes a "percentage improvement in legibility" predictor if  $F$  falls on  $L$ . The  $\Delta E$  to  $RS$  relationship described here only approximates an interval one. Yet it is clear from Figure 9 that performance prediction by this method with  $\Delta E_{std}(Y, u', v')$  is statistically most accurate at  $\Delta E_{RS_{max}}$  and increasingly conservative with respect to ensuring legible contrast as  $\Delta E$  decreases. Because Figure 9 is a plot of means, such a performance underprediction approach appears warranted.

Consider now the possibility that  $R$  does not equal  $0.87Y_M$  for some  $Y_M$ , that is, that the relative contributions to legibility of the luminance contrast and chrominance contrast unit components vary as a function of absolute levels of retinal stimulation. If it is empirically determined that the legibility contribution of chrominance contrast increases linearly with  $Y_M$ , then,

$$R = (0.87Y_M) (C_{Y_M}), \quad (20)$$

where  $C$  is the linear coefficient describing the increasing contribution of chrominance contrast as a function of  $Y_M$  and,

$$C < 1 \text{ for } (0.87Y_M) < 40, \quad (21)$$

$$C = 1 \text{ for } (0.87Y_M) = 40 \text{ and,} \quad (22)$$

$$C > 1 \text{ for } (0.87Y_M) > 40, \quad (23)$$

where ( $C = 1$  for  $(0.87Y_M) = 40$ ) has been determined for the Study 1 data.

If future evidence suggests  $C$  is non-linear with respect to some range of the  $Y_M$  continuum, then  $C$  becomes the sum of the coefficients describing the non-linear relationship as a function of  $Y_M$  over that range. This scheme provides considerable modeling flexibility. For instance,  $C$  might be an inverted "U" function of  $Y_M$ , a higher-order function, or even be described by a discontinuity, perhaps at a transition level of  $Y_M$  from predominant cone to mesopic vision.

Chrominance axes rescaling. Thus far, the  $u'$  and  $v'$  scales have been employed in the  $\Delta E$  metrics. By independently, linearly rescaling the  $u'$  and  $v'$  axes relative to one another and determining the appropriate weighting factor relative to  $Y$  (i. e., independently,

linearly rescaling Y) to achieve maximum  $r_{\Delta E}^2$ , a metric has been achieved which describes the entire data set from Study 1 including the red and magenta chromaticities (Figure 10). The results of this process are described more fully by Lippert (1984).

Figure 10 illustrates a strong relationship between combined  $\Delta E$  and  $\Delta E^2$  terms and RS (parabola  $R^2 = 0.71$ ). It is this type of fully rescaled metric with which the balance of this report will be concerned. In addition to Y, the  $\log_{10}(Y)$ , (e. g.,  $\log Y$ ) and  $L^*$  luminance scalings were considered as metric components. Figures 11 and 12 illustrate the similar RS modeling potentials of all three luminance scalings as combined with 2.2:1 rescaled ( $u':v'$ ) chromaticity axes.

It was shown earlier in this section that the convergent  $L^*, u^*, v^*$  space does not yield a  $\Delta E$  scale which describes RS well, nor does it always provide a unique color pair solution given one color specification and a required  $\Delta E$  for a desired level of legibility. However, by rescaling the  $L^*$ ,  $u^*$ , and  $v^*$  axes to fit the Study 1 data as well as possible, an RS description similar to those of the non-convergent spaces is obtained. This  $\Delta E$  to RS relationship is shown by Figure 13, which appears very similar to Figures 10-12. It is important to note that the convergent nature of  $L^*, u^*, v^*$  space is not completely eliminated by the

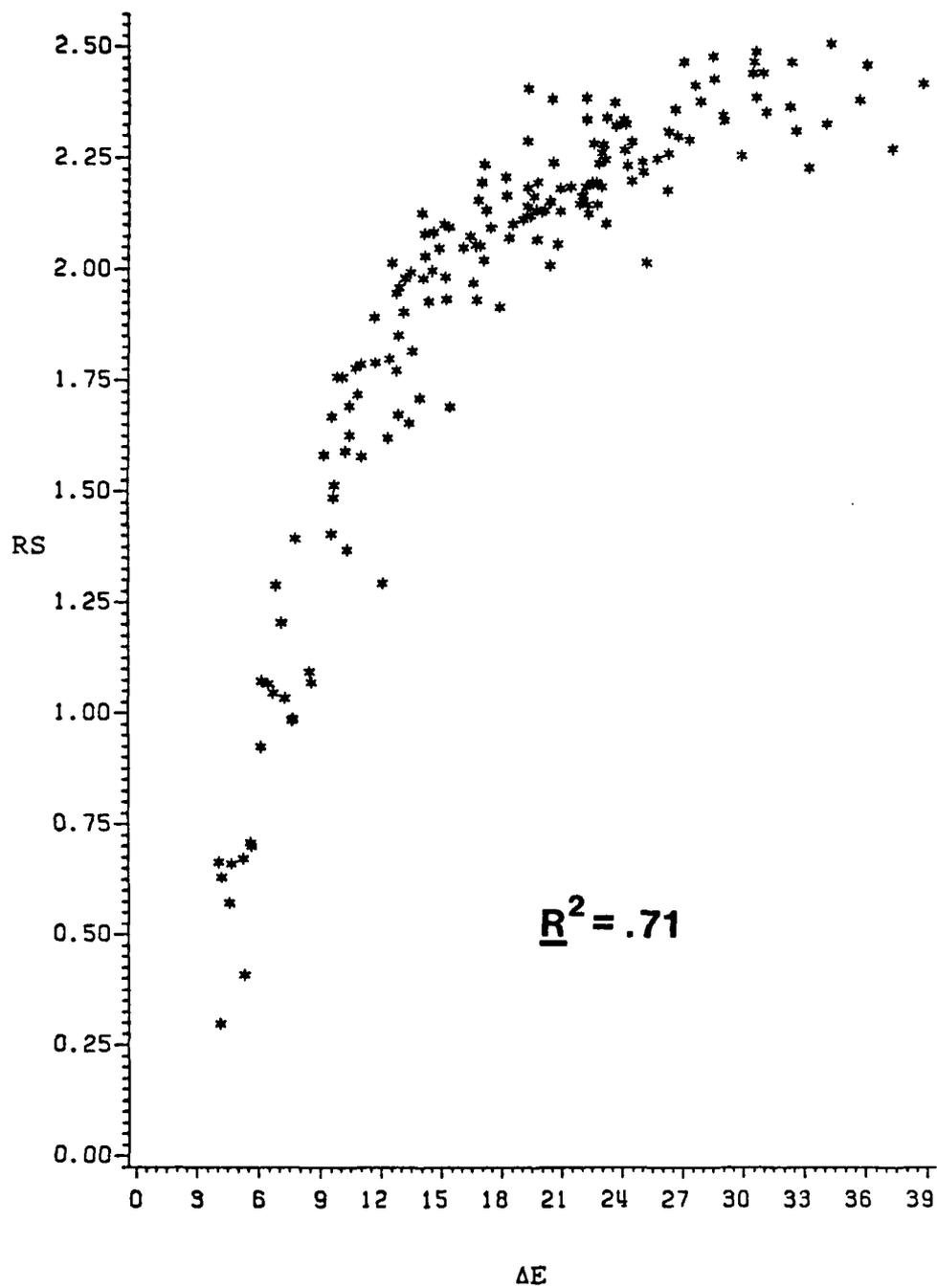


Figure 10.  $\Delta E(Y, u', v')$  by RS scattergram: rescaled chrominance axes, Study 1 (Lippert, 1984).

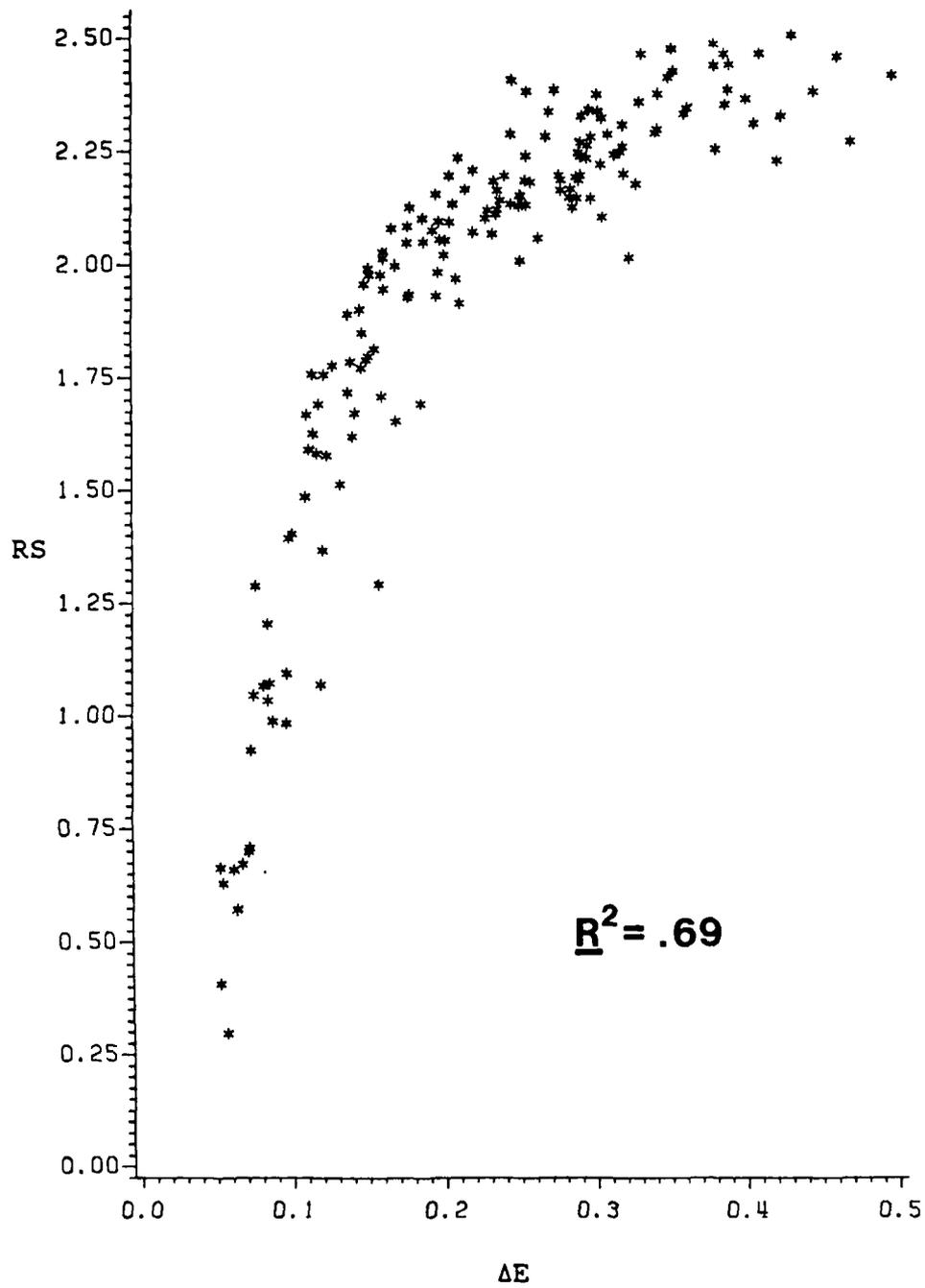


Figure 11.  $\Delta E(\text{LogY}, u', v')$  by RS scattergram:  
 rescaled chrominance axes,  
 Study 1 (Lippert, 1984).

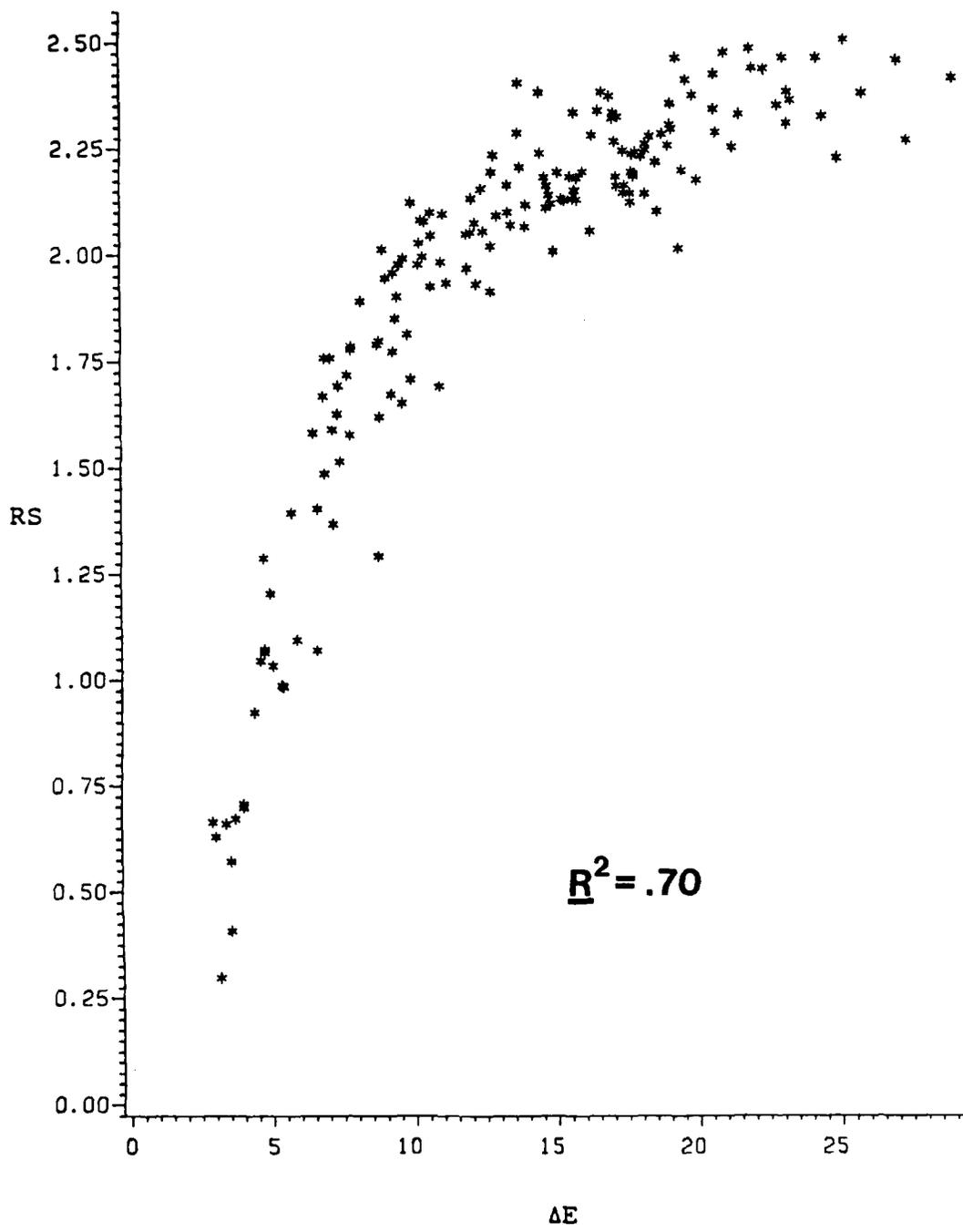


Figure 12.  $\Delta E(L^*, u', v')$  by RS scattergram:  
 rescaled chrominance axes,  
 Study 1 (Lippert, 1984).

rescaling procedure. Therefore, it will in some cases provide similar, although not unique, solutions.

Table 2, RESULTS, contains a listing of  $\Delta E$  component weightings and regression coefficients and Table 3, RESULTS, lists the  $\Delta E$  component weightings and the luminance-generalized,  $\Delta E$ -standardized expressions for the  $Y, u', v'$ ,  $\log Y, u', v'$ ,  $L^*, u', v'$ , and  $L^*, u^*, v^*$  metrics.

Luminance-generalization and  $\Delta E$ -standardization of the  $(Y, u', v')$  metric with independently, linearly rescaled axes is based on the complete Study 1 stimulus/response set as follows:

$$Y_M = 46.6 \text{ cd/m}^2 \text{ (averaged across all Ts),} \quad (24)$$

$$\text{weighting}_{u'} = 110, \text{ and} \quad (25)$$

$$\text{weighting}_{v'} = 50. \quad (26)$$

Factoring 50 from the weightings,

$$\text{weighting}_{u'} = 50 ( 2.2 u' ), \text{ and} \quad (27)$$

$$\text{weighting}_{v'} = 50 ( 1.0 v' ); \text{ therefore,} \quad (28)$$

$$\text{weighting} = 50 \quad (29)$$

for both chrominance components (2.2 remains with  $u'$ ).

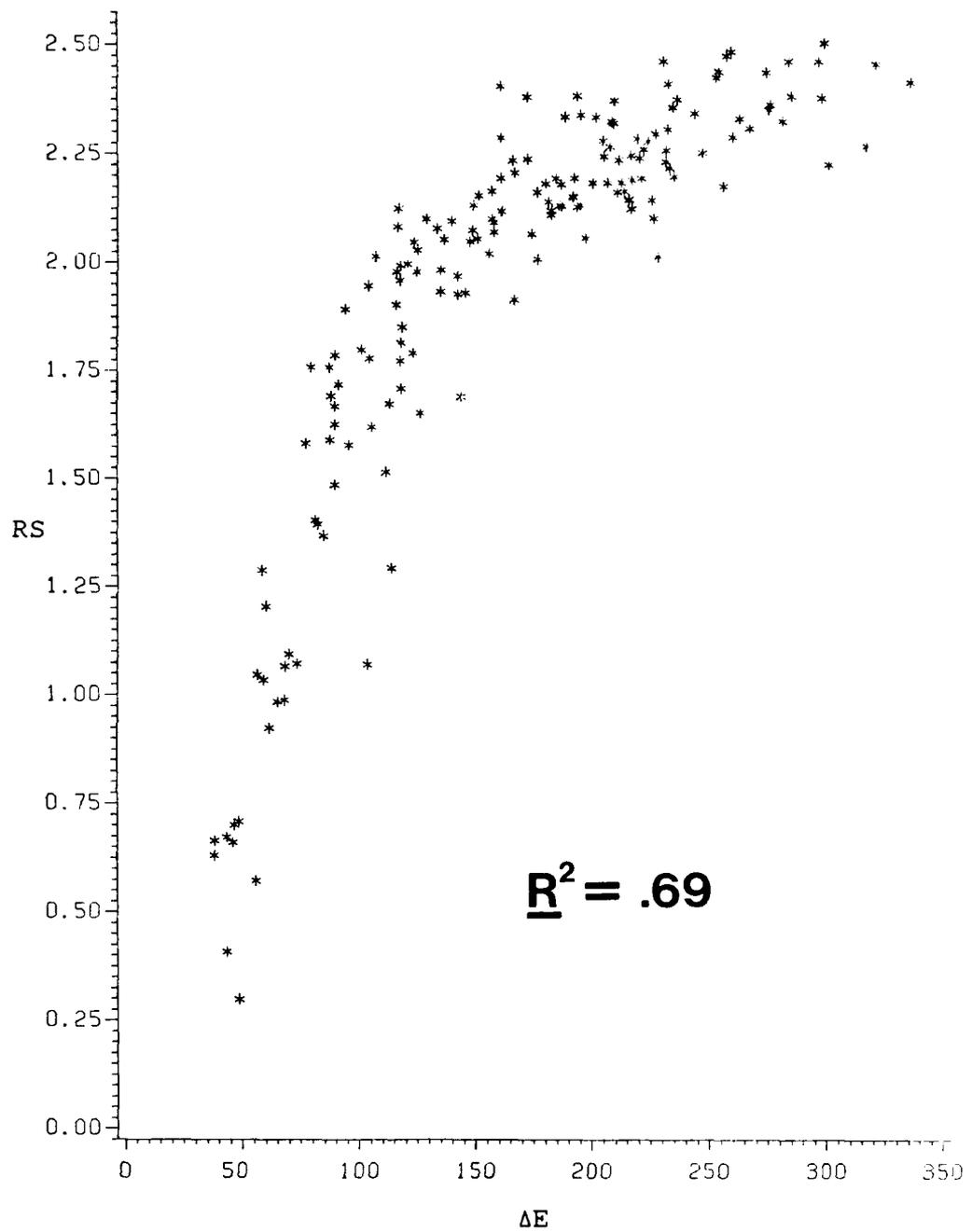


Figure 13.  $\Delta E(L^*, u^*, v^*)$  by RS scattergram:  
rescaled axes, Study 1 (Lippert, 1984).

Then,

$$R = \text{weighting} / Y_M = 50 / 46.6 = 1.073 \text{ and,} \quad (30)$$

$$W = 1.073Y_M. \quad (31)$$

Now, from Figure 10,

$$\Delta E_{RS_{\max}} = 30, \quad (32)$$

corresponding to an  $RS_{\max} = 2.37$  compared to 2.17 for the metric excluding red and magenta. The red and magenta inclusive metric requires greater predicted RS for 100% legibility and is therefore significantly more conservative. Now, the relationship between the weighting factor and  $\Delta E_{RS_{\max}}$  is,

$$50 / 30 = 1.6667. \quad (33)$$

Selecting  $\Delta E_{RS_{\max}} = 100$ ,

$$W_{\text{std.}} = (100) (1.667) = 166.7, \text{ and} \quad (34)$$

$$S = W_{\text{std.}} / W = 166.7 / 1.073Y_M = 155.3. \quad (35)$$

Completing the process via (18),

$$\Delta E = (((155 / Y_M) \Delta Y)^2 + (367\Delta u')^2 + (167\Delta v')^2)^{0.5}. \quad (36)$$

For a logY metric, the weightings are 1.43 ( $\Delta u'$ ) and 0.65( $\Delta v'$ ), resulting in the final expression,

$$\Delta E = [((250/\log Y_M)(\log(Y_T/Y_B))^2 + (7.67 \Delta u')^2 + (3.49 \Delta v')^2]^{0.5}, \quad (37)$$

for the L\* version, the weightings are: 79( $\Delta u'$ ) and 36( $\Delta v'$ ), resulting in,

$$\Delta E = [((260/L^*_M) \Delta L^*)^2 + (440 \Delta u')^2 + (200 \Delta v')^2]^{0.5}. \quad (38)$$

The weightings for L\*,u\*,v\* are 1.0 ( $\Delta L^*$ ), 0.00583( $\Delta u^*$ ), and 0.00278( $\Delta v^*$ ), resulting in

$$\Delta E = [((595/L^*_M) \Delta L^*)^2 + (0.75 \Delta u^*)^2 + (0.36 \Delta v^*)^2]^{0.5} \quad (39)$$

Equations (36), (37), (38), and (39) are the metrics of particular interest to this report. They do not include any task effect other than that of color contrast and are designed for use as relative indices of legibility.

Figure 14 is a scattergram of luminance-generalized,  $\Delta E$ -standardized  $\Delta E(Y,u',v')$  versus RS from Study 1. Except for an insignificant amount of numerical disagreement due to the different  $\Delta E$  computational procedures used for Figure 14

versus Figure 10, the standardized formula provides the same  $\Delta E$ -to-RS relationship as does the non-standardized version. The  $\Delta E$  scale in Figure 14 relates  $RS_{\max}$  to  $\Delta E = 100$  and, if used as a percent improvement in legibility scale, provides underestimations of the observed RS means below  $RS_{\max}$ .

### Research Objectives

Study 1 determined that a non-convergent  $\Delta E$  space is appropriate for RS modeling. It provided a large data set with which to develop  $\Delta E$  metrics but tested only positive presentation polarities and only at a single  $Y_M = 46.6 \text{ cd/m}^2$  (averaged) across Ts. Study 2 questions the validity of the assumption that the effect of  $\Delta E$  on RS is independent of the absolute level of retinal stimulation, that is, that R is a parametric constant, by employing a  $Y_T = 20 \text{ cd/m}^2$ , and considers the application of any single metric to both positive and negative presentation polarities.

The luminance-generalized,  $\Delta E$ -standardized metrics are employed in Study 2 to determine the merit of the assumption that R is constant. To the extent that the metrics for the positive presentation polarities in Study 1 adequately describe RS in Study 2, the method of legibility prediction at low contrast levels using  $\Delta E$  would be confirmed and generalized.

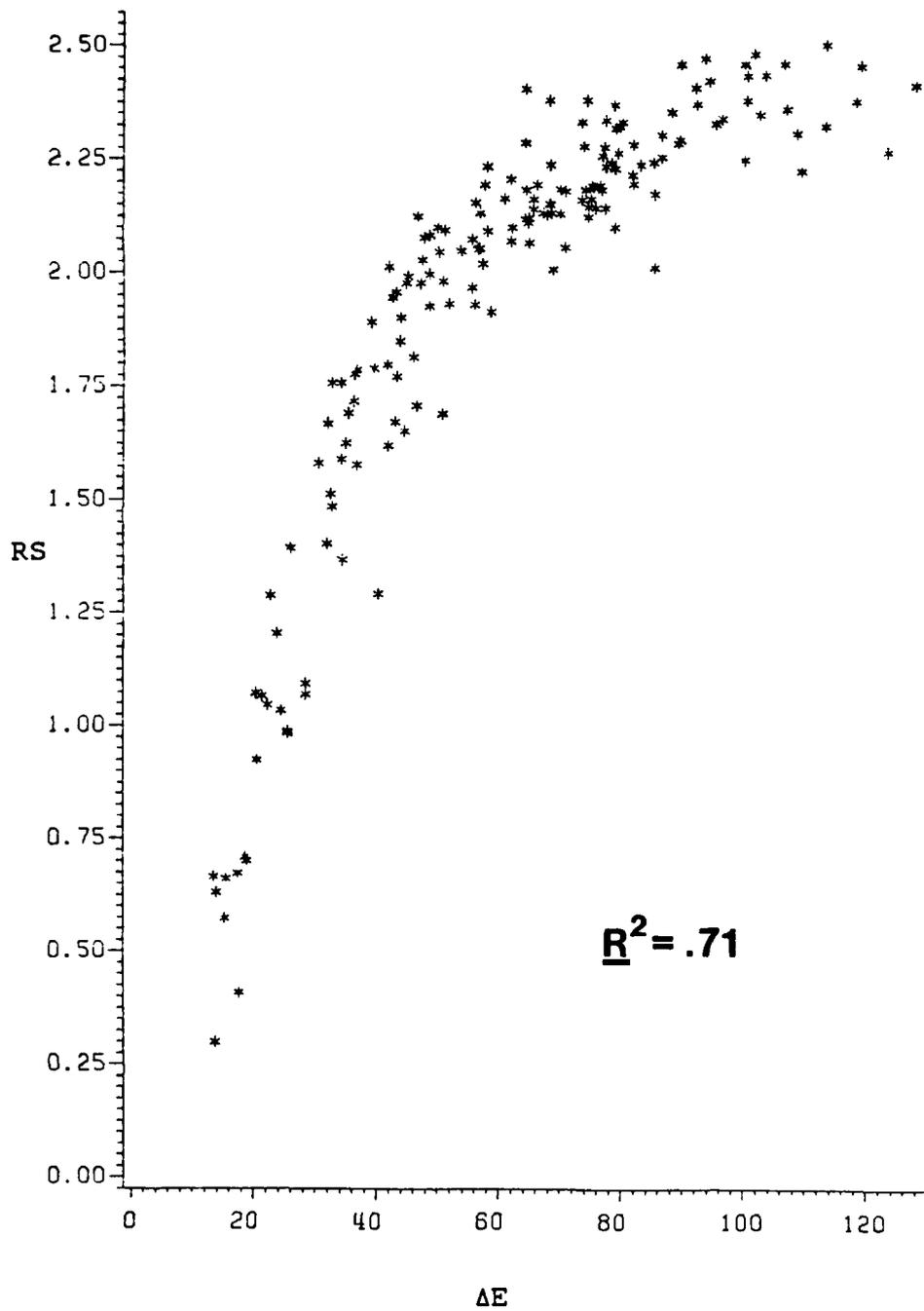


Figure 14.  $\Delta E$  in luminance-generalized,  $\Delta E$ -standardized  $(Y, u', v')$  space by RS scattergram, Study 1 (Lippert, 1984).

## METHOD

### Digital Color Video System

The display/radiometric system used in this research is actually an evolved combination of computer, refresh-type digital image processor, CRT, radiometric measurement, and laboratory interconnection systems. It is capable of displaying color images within the design limits of its components, scanning images spectroradiometrically, and reducing the resultant data to photometric parameters. The measurement subsystem is calibrated to a National Bureau of Standards-traceable spectral radiance source (approximately standard illuminant A), allowing the user to develop, display, verify, and report the power spectra of color images. A technical report describing the system's functional requirements has been published (Farley and Gutmann, 1980). This subsection addresses the preparation of the laboratory system for this particular experimentation.

Monitor. An Aydin Model 8025 19-inch diagonal high-resolution color monitor serves as the display. It employs a delta-gun (R, G, and B) Mitsubishi shadow mask CRT incorporating P22 phosphors and a 0.29 mm triad pitch. The raster is 2:1 positively interlaced and the normal aspect

ratio is adjusted to 1:1 (square active display). The monitor is driven by an IIS Model 70 digital refresh image processor (RS-170 standard), which delivers 512 x 512 picture element (pixel) images to each of the CRT guns for the square trichromatic display. Refresh rate is 60 fields/s (30 frames/s).

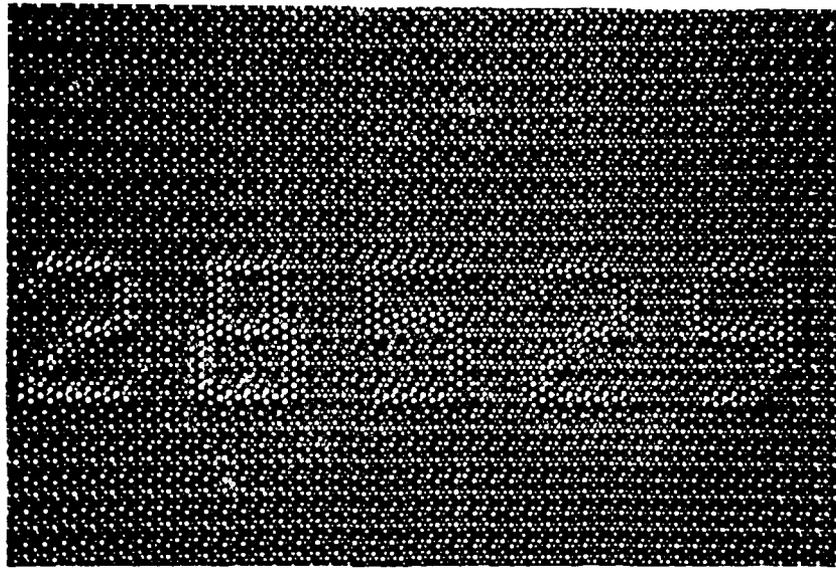
CRT convergence. Visual inspection of variously colored numerals and backgrounds indicates that normal CRT dynamic convergence adjustment procedures using a cross-hatch pattern are suboptimal. The slightest misconvergence results in visible dark outlines for the numerals, outlines which are unwanted display artifacts that increase perceived numeral/background contrast beyond that produced by the photometrically controlled stimulus parameters. It was decided to adjust the dynamic convergence at each experimental task position of the screen while actual experimental images were being displayed. The procedure was to adjust for minimum perceived contrast, which is obtained with perfect convergence, but which is not obtainable on the monitor employed over the entire area of any task location, much less over the entire screen. For this reason as well as the spectral radiance nonuniformities discussed below, the accuracy of spatially-dependent displayed color specifications remains somewhat in question. Observers noted that the numerals are less easily read given the above

adjustment procedure, especially at low contrast levels. Black and white facsimiles of best and poor attainable convergence for red numerals on an achromatic background are shown in Figure 15, a and b, respectively. The "poor" example is within the manufacturer's specifications, indicating that a global adjustment strategy is not sufficient when local convergence is critical.

Raster delay. A response trigger was used by subjects to begin and end experimental trials while automatically recording response time as measured by a 1-ms resolution clock. A lag time was specified for the display of task numerals once a trial was initiated. For task 1 (e.g., top center 3-digits) lag time is negligible. For task 2 (e.g., left center 4-digits) and task 3 (e.g., right center 5-digits) it is 8 ms. This measure of lag time assumes the clock and first HUD field begin at the end of the raster's vertical retrace following trigger depression and is linearly related to the number of lines below raster-line 1 of the numerals' bottom line (first field). When the trigger is released, the clock is stopped and an inter-stimulus image (ISI) is displayed. End-of-trial functions occur within approximately 1 ms of subject response.

However, there is an inaccuracy inherent in this procedure. It is due to the interval between the final refresh of the numerals (ending a trial) and the first paint

a



b

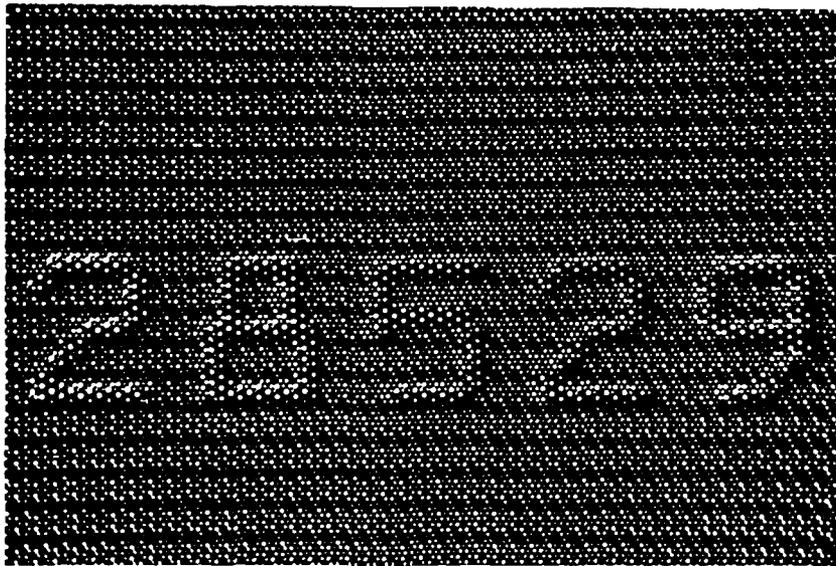


Figure 15. Black and white photographic examples of dynamic convergence of the shadowmask display.

of the ISI over the task numeral location. The P22 phosphors' persistences range from about 22 microseconds (B) to 1 ms (R), presenting no significant problem, but the retinal persistence of the task numeral images may be great enough to last until re-stimulation by the ISI on the following raster. This would introduce a systematic increase in the time the numerals are visible beyond that recorded by the clock. It is a concern that response correctness might be inflated, because the observer could continue to process task information after the end-trial response. Because subjects were instructed to perform as quickly and accurately as possible and because it is assumed they will not end the trial before extracting the requested information, the response time measure, which is accurate to within approximately 2 ms once corrected through subtraction of the fixed image onset delay of 8 ms (Tasks 2 and 3 only) via software, is considered of primary importance.

Radiometric measurement system. A subsystem very critical to experimental control is the radiometric measurement system (MS) described in block diagram form by Figure 16. The MS is a modified Gamma Scientific, Inc., visible band system. To determine the power spectrum of a source of light, some of its emission is collected by a fibre-optic cable and fed to a monochromator which optically processes that energy into a spectrum of known location in

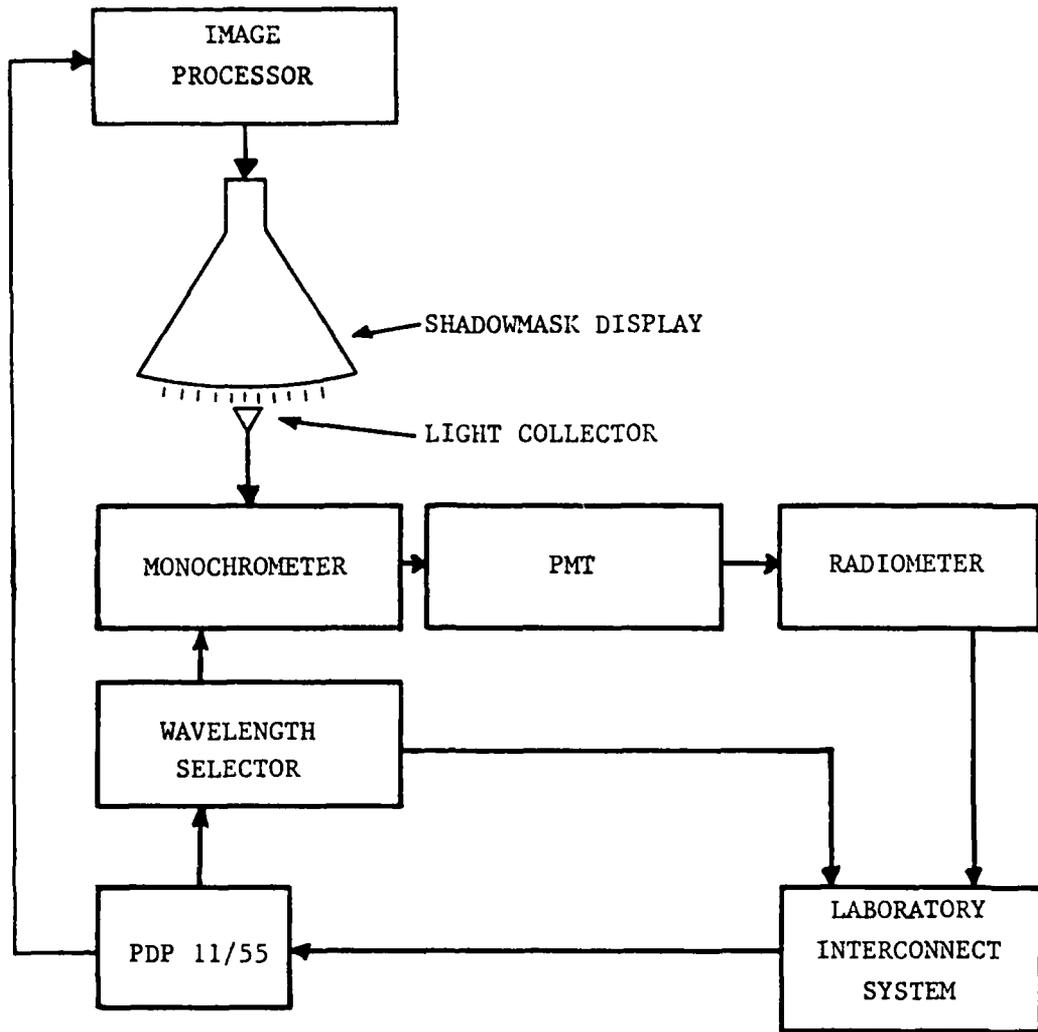


Figure 16. Digital color video system block diagram.

space. The spectrum is sampled (scanned) over a specified range (380 to 760 nm in 5.0-nm increments for this study) and the narrow bandwidth samples (half-power bandpass of the monochromator is 5.0 nm) are amplified by a photo-multiplier tube (PMT). The sample values are stored in computer files (scan files) which contain radiance values as a function of wavelength. Photometric values are computed directly from the radiance files via software.

Calibration. It is necessary to calibrate the MS relative to a full-spectrum standard light source, the power spectrum of which is known and is stored in the computer. The amplified radiance samples collected from this source are sent to the computer in the form of analog voltages which are digitized. Because each sample's wavelength is known, its corresponding absolute radiance value, stored in the computer, is simply divided by the sample value to create correcting, or scaling, factors. Scaling factors permit the accurate characterization of unknown spectral distributions even though the gain and linearity of the MS (although not the absolute output of the radiance source) are expected to vary with time.

A source of unknown spectral distribution (a CRT image) is sampled at the same wavelengths as is the standard light source. The unknown's sampled digital values are multiplied by the corresponding scaling factors, yielding a relative

spectroradiometric characterization. This scanfile may be weighted by the 1931 C.I.E. 2-degree color-matching functions and integrated across the spectrum to yield tristimulus values and, therefore, any desired photometric transforms.

A new spectral radiance standard utilizing a General Electric DYT 19-volt rated quartz tungsten-halogen projection lamp has been developed in-house. A 90-min repeated-scan procedure indicated the combined radiance source/MS drift to be on the order of 0.5 % in luminance and 0.0002 in either CIE x or y.

The radiometric measurement techniques employed assume that a uniform light source is being scanned. For the CRT, radiometric measurements exhibit no appreciable change as a function of screen-to-collector distance or incidence angle (within reasonable limits). In other words, the shadow mask display behaves much like a cosine radiator and is appropriately measured using our current procedure. The DYT radiance source was calibrated by Hoffman Engineering Corporation by sampling only the central 1.27 cm spot on its diffuser. This procedure is matched and maintained for all MS calibrations.

Sources of error. There are two major sources of error associated with the display/radiometric measurement system. The first is the MS itself. The PMT is particularly prone to signal-to-noise and drift problems, which are controlled by averaging multiple samples at each wavelength and frequent MS calibration.

The second source of error, the output stability of the monitor, can, in general, only be recorded for post-experiment adjustment of stimulus parameter specifications, or it may simply be accepted as a source of noise in experiment response data. The reported color parameters for this study were determined at the midpoint of the data collection sequence. R, G, and B command values were not adjusted during data collection sessions.

Monitor characterization requires that each gun (R, G, and B) be scanned throughout its available luminance range using the calibrated MS. This results in bits-to-luminance (computer command-to-display output) functions for each gun. A computerized algorithm searches look-up tables of 1931 CIE tristimulus values derived from the sampled points along each function which are additively combined and tested to provide R, G, and B bit-values for any desired CRT-attainable stimulus color.

System stability. As explained earlier, the luminance and chrominance of displayed colors should be controlled to within 1 JND. Repeated pre-experiment scans of the monitor indicate an overall measurement/display system stability. The luminance instability was determined through nightly scanning of all experimental stimulus colors over the 8-day data collection period and equalled less than 1.5 % of any of the stimulus luminance specifications in Table 4, Appendix A.

System baseline tests. A set of baseline tests was developed to evaluate the electrical and mechanical integrity of the display/radiometric system and provide a means of detecting changes in system component performance. These tests were performed before monitor characterization for this study.

Experiment control. The laboratory's PDP 11/55 minicomputer and associated networks were used for real-time control of all display/radiometric measurement system operations for this experiment, including data collection and storage. Data were transferred to the Virginia Tech IBM 370 for analyses using the Statistical Analysis System.

## Stimuli

Active display. All stimuli were presented on the Aydin monitor in a central 26 cm x 26 cm area defining the background display. The 512 x 512 pixel image thus contains pixels equally spaced horizontally and vertically, facilitating programming of the 7 x 9 dot-matrix Huddleston font numerals which were designed for maximum legibility in high ambient illumination environments (Snyder, 1980). At a viewing distance of 0.76 m, the active display subtended 20 visual degrees horizontally and vertically, resulting in a limiting display resolution of approximately 13 cycles/degree. The 0.76 m viewing distance in Study 2 was increased from the 0.5 m viewing distance in Study 1. This was done to determine whether this limited range of viewing distance had a marked effect on the ability of Study 1 metrics to describe RS in Study 2. Because visual accommodation to the displayed imagery and the visual angle subtended at the eye of a numeral, or a matrix dot, or a numeral string, or the active display area, all change with viewing distance, none of these factors could be subjected to analysis. The numerals (one pixel per dot) subtended 16 horizontal and 21 vertical visual arcminutes. The flight control HUD was embedded into the background (Figure 17) such that HUD and background photometric parameters are independent.

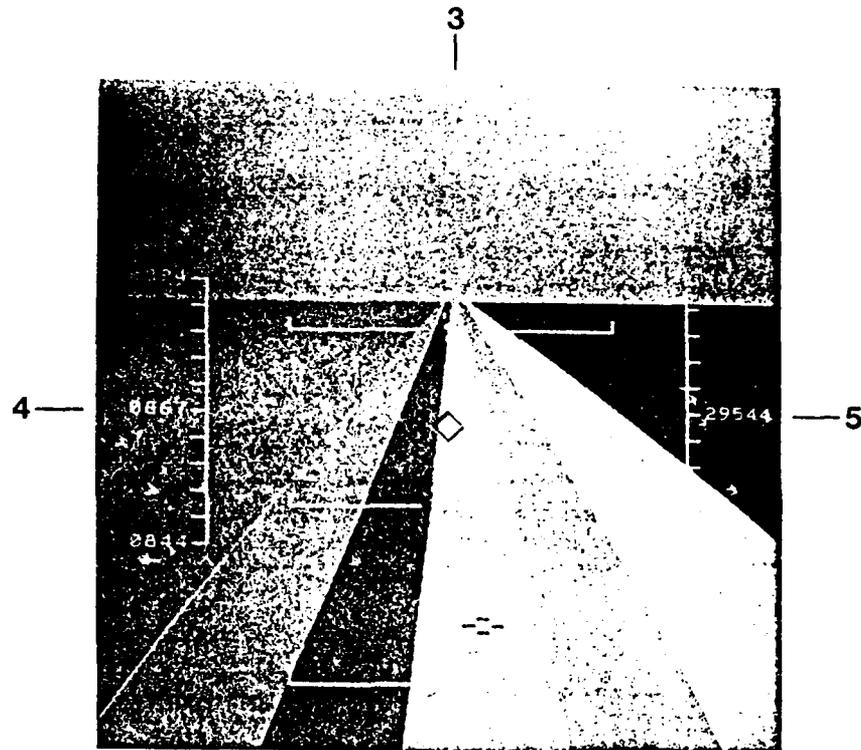


Figure 17. Black-and-white photographic example of HUD symbology, background fields, and the 3-, 4-, and 5-digit reading tasks.

Numeral strings and background. On a given trial color contrast was accomplished by displaying a HUD and background of known color parameters. The task numerals are completely surrounded by a background field of desired parameters. While each of the 10 background fields is independently addressed, only three of the fields surround task numerals. Other fields were of color parameters randomly selected from the entire pool of experimental background colors.

The nominal HUD and background chromaticities selected for Study 2 are depicted on the CIE 1931 (x,y) chromaticity diagram by Figure 18, and the measured stimulus color parameters are listed in Table 5, Appendix A. Included are achromatic (ACH), yellow-green (Y-G), and red (RED) Ts and purple (p1,p2), magenta (m1,m2), red (r1,r2), green (g), and achromatic (a) Bs. The vertices of the triangle inset represent the chromaticities of the Mitsubishi CRT P22 phosphors.

Interstimulus image (ISI). A  $2^2$  pixel black-and-white checkerboard filled the active display area between trials. The white squares were displayed at the constant experimental numeral luminance ( $20 \text{ cd/m}^2$ ). The intended functions of the ISI were to maintain the luminance adaptation state of subjects between trials, minimize successive chromatic adaptation confounds (Wyszecki and Stiles, 1967, p. 235), provide interference with any retinal

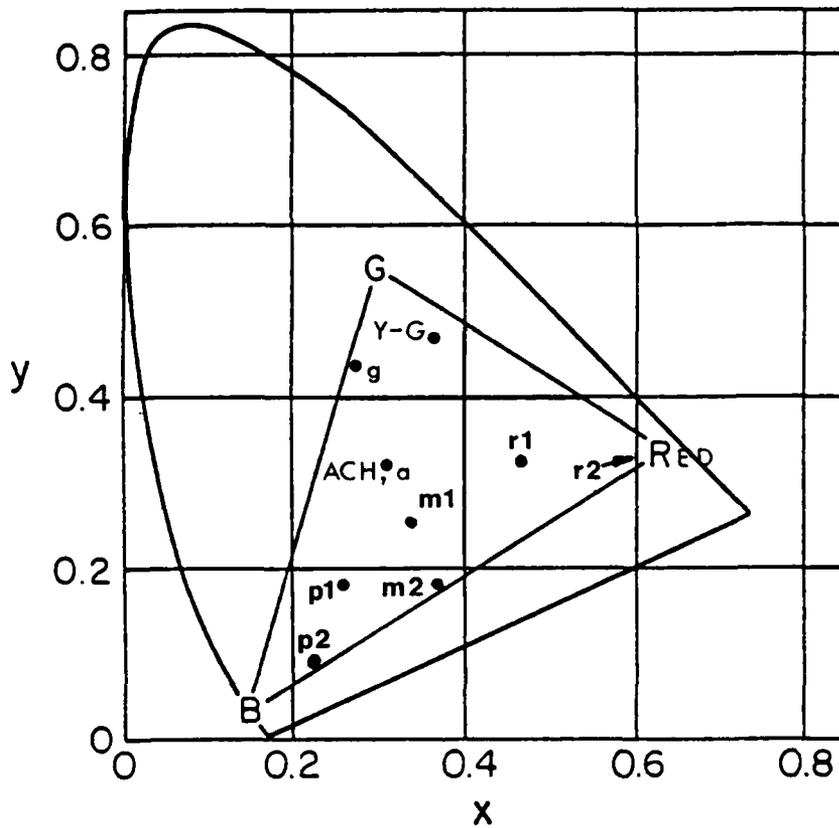


Figure 18. Chromaticities for HUD symbology and background fields on the 1931 CIE Chromaticity Diagram.

persistence of the task numeral images, and maximize the trial-to-trial luminance output stability of the monitor.

### Experimental Design

The study investigated reading speed (RS) as a function of color contrast operationally defined in terms of (1) numeral, or target (T) chromaticity, (2) background (B) chromaticity, and (3) target-to-background (T/B) luminance modulation (M).

Luminance and chrominance contrast. The attainability of desired color parameters is limited by the phosphors and operating characteristics of the CRT and the digital resolution of the image processor employed, and varies from maximum flexibility at low purities to greatly reduced flexibility at high purities. Parameter maintainability is an expression of the display system stability, as specified previously. Scan-to-scan variability in stimulus L measures indicated a maintainability of  $\pm 2\%$  over a two week period. For this reason, no problem existed in maintaining T or B luminances to achieve linear steps in M of approximately 0.09.

Experimental manipulations. The experimental design consisted of seven levels of M combined factorially with eight levels of B, resulting in 56 color contrast (T/B)

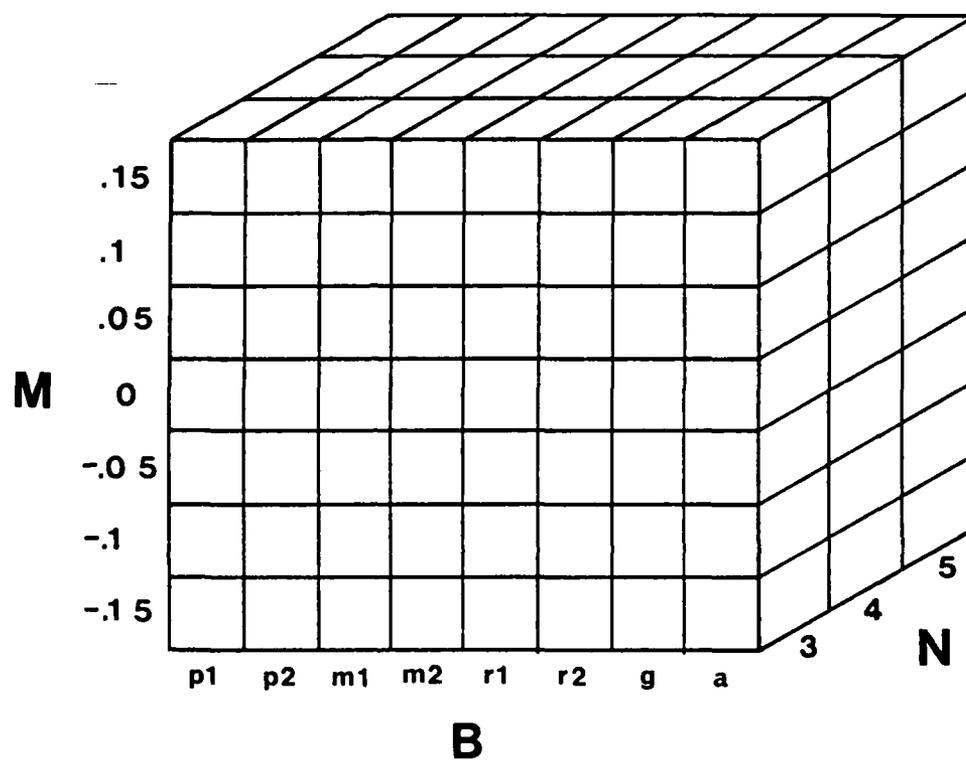


Figure 19. Nominal experimental manipulations of numeral-to-background luminance modulation (M), background chromaticity (B), and number of reading task digits (N) for each of achromatic, yellow-green, and red numeral chromaticities (T).

combinations for each of the three Ts in each of three reading tasks, 3, 4, and 5 digits (N). Figure 19 shows the nominal B x M x N stimulus matrix for each T. Note that luminance modulation may take on positive or negative values as defined for Study 2 ( $M = (L_T - L_B)/(L_T + L_B)$ ), corresponding to positive presentation polarities (i. e., lighter characters on a darker background) and negative polarities, respectively. The Study 1 experimental design matrix differs from that for Study 2 only in that different levels of the T, B, and M dimensions were employed (Table 4, Appendix A). As stated previously, Study 1 concentrated on positive presentation polarity between  $46 \text{ cd/m}^2$  numerals and a parametric selection of chromaticities while Study 2 used  $20 \text{ cd/m}^2$  numerals in both positive and negative presentation polarities with a predominance of chromaticities from the red, magenta, and purple chromaticity sectors. In both studies, the white T (ACH) had only 55 T/B combinations because it alone was nearly identical chromatically to the achromatic B. The lowest  $M_{B_a}$  stimulus condition for  $T_{ACH}$  was therefore eliminated because there was essentially no luminance or chrominance contrast to distinguish T from B.

The experiment consisted of four replications of the T x B x M x N design for each of six subjects, totalling 12,024 trials. The first replication was used as a training exercise while replications 2-4 constituted the 9018

experimental observations acquired for the present analysis. Total trial time averaged about 10 s/trial, resulting in 33 hours of data collection.

### Subjects

Six volunteers (three females) from the Virginia Polytechnic Institute and State University student pool served as subjects. All were screened for normal color vision using Dvorine Pseudo-Isochromatic Color Plates (there were no misses) and for 20/20 uncorrected near and far visual acuity using a Bausch & Lomb Orthorater. Upon completion of all experimental sessions, the experimenter debriefed each subject, offered a summary of the results at a future date, and paid each subject 20 dollars.

### Procedure

Screened subjects participated in one 60-min session every day (completing one replication per session). The instructions, to read as quickly and accurately as possible one specified numeral string per trial, were read to each subject at the first session. At each session, a subject was presented the ISI for three to five minutes, allowing his/her visual system to adapt to the moderate luminance of the display. The sessions were divided into three subsessions, each equalling one-third of one replication of the study, with a short rest period between subsessions.

Each subsession consisted of 167 trials in three blocks. Each block was of one T, and the presentation order of Ts was randomized with the constraint that each of the three T-blocks began one subsession in each session. Within blocks, the B x M x N combinations were randomized with the constraint that a complete replication be presented within one session.

Figure 20 is a photograph of a subject seated in typical posture at the subject station. A forehead rest maintained a centered active-field subtense of 20 degrees. Not shown is a black surround, subtending approximately 55 vertical and 80 horizontal visual degrees, which was fitted over the active display to mask the monitor casing. The experiment was conducted without room illumination to control for possible chromatic adaptation effects of the experiment room walls. A lightly spring loaded momentary solid state switch set into a Dan Wesson finger-grooved pistol grip was used by subjects as an index finger trigger for the display. The grip was mounted for comfortable usage with a 15-degree forward tilt and a removable wooden block for resting the heel of the hand if desired. This assembly is free-standing, may be positioned by subjects for maximum ease of use, and could be used with either hand. Subjects typically leaned into the forehead rest as in Figure 20, but a padded lumbar support was provided for those who wished to sit

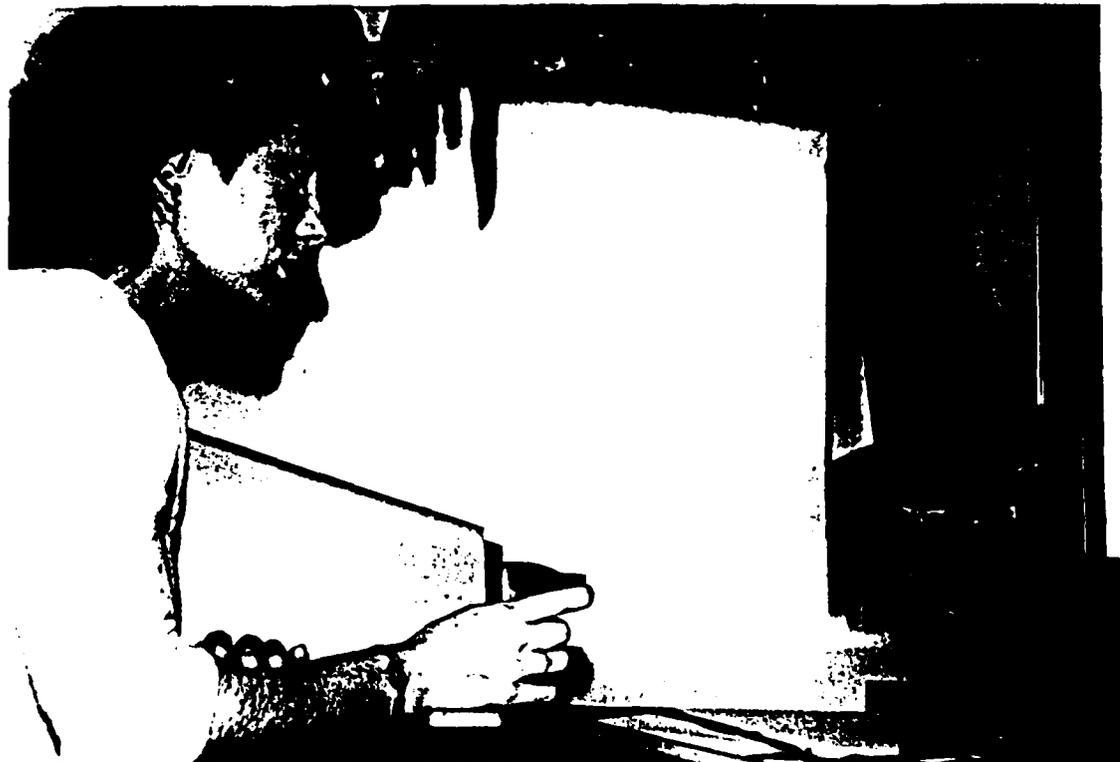


Figure 20. Subject station.

upright. The padded seat was adjusted in height for each subject to match eye-height to center-screen height. A foot-rest was furnished for one subject requiring it. The experimenter controlled each session seated at a VDT behind the subject station.

Each trial began when the experimenter informed the subject of the reading task for that trial (i. e., airspeed, heading, or altitude). The experimenter received this information from the VDT. When ready, the subject initiated the trial by triggering the solid state switch and holding it in. When the requested information was extracted from the display, the subject released the trigger, stopping the response-time clock (automatically recording and indexing elapsed time for the trial) and replacing the experimental field with the ISI. The subject then verbalized the digits read for the experimenter, who checked the verbalization against the correct numeral sequence and kept a cumulative error count by hand. Error rate in Study 2 is as in Study 1 and in Post's (1983) experiment, about 2 percent. For correct responses, the experimenter struck the return key or, some number of zeros (erroneous digits) was entered followed by the return key, which always signaled the computer to begin the next trial. The computer then prompted the experimenter who in turn prompted the subject for the next trial, etc. Care was taken not to force any

subject into a minimum trial-time mental set. Rather, the emphasis was on relaxed attention to the display for subject initiation of each trial.

## RESULTS

The present study (2) was conducted to determine the generalizability of any metrics developed from the Study 1 data. The success of four Study 1 metrics in describing the Study 1 data assists researchers of this type of performance because it precludes the analysis of variance procedures used in Study 1 to determine levels of performance associated with levels of color contrast. It has been demonstrated that an appropriate continuous scale of  $\Delta E$  may be applied to the description of RS. The analysis of Study 2 therefore reduces to least-squares regressions of  $\Delta E$  from Study 1 metrics with RS from Study 2.

CIE 1931 (X,Y,Z) tristimulus values, 1976 (u',v') coordinates,  $L^*$ , and  $\log_{10}(Y)$  were computed for each T and B. The 167 stimulus color combinations for each study were transformed into several  $\Delta E$  scales based on the above luminance components, each combined with rescaled (e. g., 2.2:1 u':v') chromaticity coordinates as described in the INTRODUCTION. In addition,  $\Delta E(L^*,u^*,v^*)$  was computed. These continuous one-dimensional variable scales were evaluated as predictors of RS via regression analysis.

The regression coefficients of the rescaled Study 1 metrics are summarized in Table 2 for the RS data sets from Study 1 and Study 2. The Statistical Analysis System GLM

Table 2.  $\Delta E$  Component Weightings and Regression  
Coefficients for the Rescaled Metrics

----- RS Data Set -----				
Parameter	Study 1		Study 2	
	b-Coefficient*	R <sup>2</sup>	b-coefficient*	R <sup>2</sup>
$\Delta E(Y, u', v')$ :		.923		.951
Intercept	1.7087		1.6522	
N	-0.3980		-0.3655	
$\Delta E$	0.1598		0.1631	
$\Delta E^2$	-0.0028		-0.0025	
$\Delta E(\log Y, u', v')$ :		.904		.929
Intercept	1.8205		1.7081	
N	-0.3980		-0.3655	
$\Delta E$	12.1732		12.5470	
$\Delta E^2$	-17.0113		-16.6709	
$\Delta E(L^*, u', v')$ :		.914		.946
Intercept	1.7923		1.7343	
N	-0.3980		-0.3655	
$\Delta E$	0.2087		0.1909	
$\Delta E^2$	-0.0050		-0.0047	

Table 2, continued.

---

$\Delta E(L^*, u^*, v^*)$ :		.909		.940
Intercept	1.8581		1.8022	
N	-0.3980		-0.3655	
$\Delta E$	0.1217		0.1206	
$\Delta E^2$	-0.0019		-0.0017	

---

\* all  $t > 3$ ,  $p < 0.003$

procedure Student's t-tests on each b-coefficient determined each slope to be non-zero (i. e., all of the model terms are meaningful predictors of RS  $p < 0.003$ ). The effects of N and  $\Delta E$  are independent in Study 2 as they are in Study 1.

The Table 2 regression models are of the form,

$$RS = x_0 + x_1N + x_2\Delta E + x_3\Delta E^2 + e, \quad (40)$$

where  $x_0$  is the RS intercept,  $x_1$  is the b-coefficient (a negative value) for the number of digits read,  $x_2$  is the b-coefficient for the root sum of squares of the independently, linearly rescaled luminance and chrominance color contrast components ( $\Delta E$ ),  $x_3$  is the b-coefficient (a negative value) for  $\Delta E^2$ , and  $e$  is random error variability.

Figures 21 and 22 are scattergrams of generalized, standardized  $\Delta E(Y,u',v')$  by RS and  $\Delta E(L^*,u^*,v^*)$  by RS in Study 2, respectively. The RS means for positive and negative presentation polarity conditions are interspersed on the figures, showing no spread of the scatter (nor reduction in  $R^2$ ) when either metric is applied to both polarities.

Working directly from these metrics, 100% legibility is predicted to require a displayed luminance ratio of 2.8, or

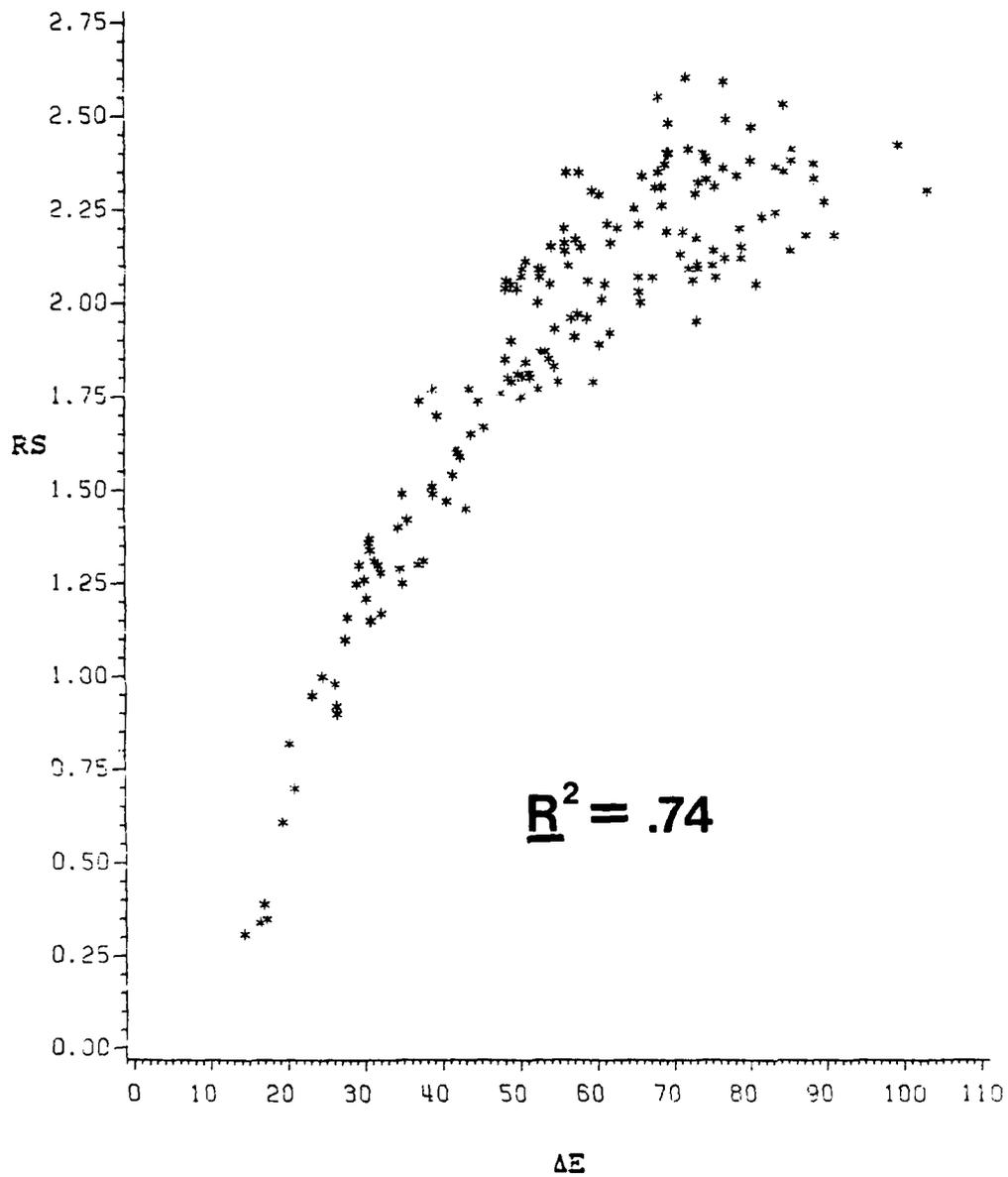


Figure 21.  $\Delta E$  in luminance-generalized,  $\Delta E$ -standardized  $(Y, u', v')$  space by RS scattergram, Study 2 (Lippert, 1985).

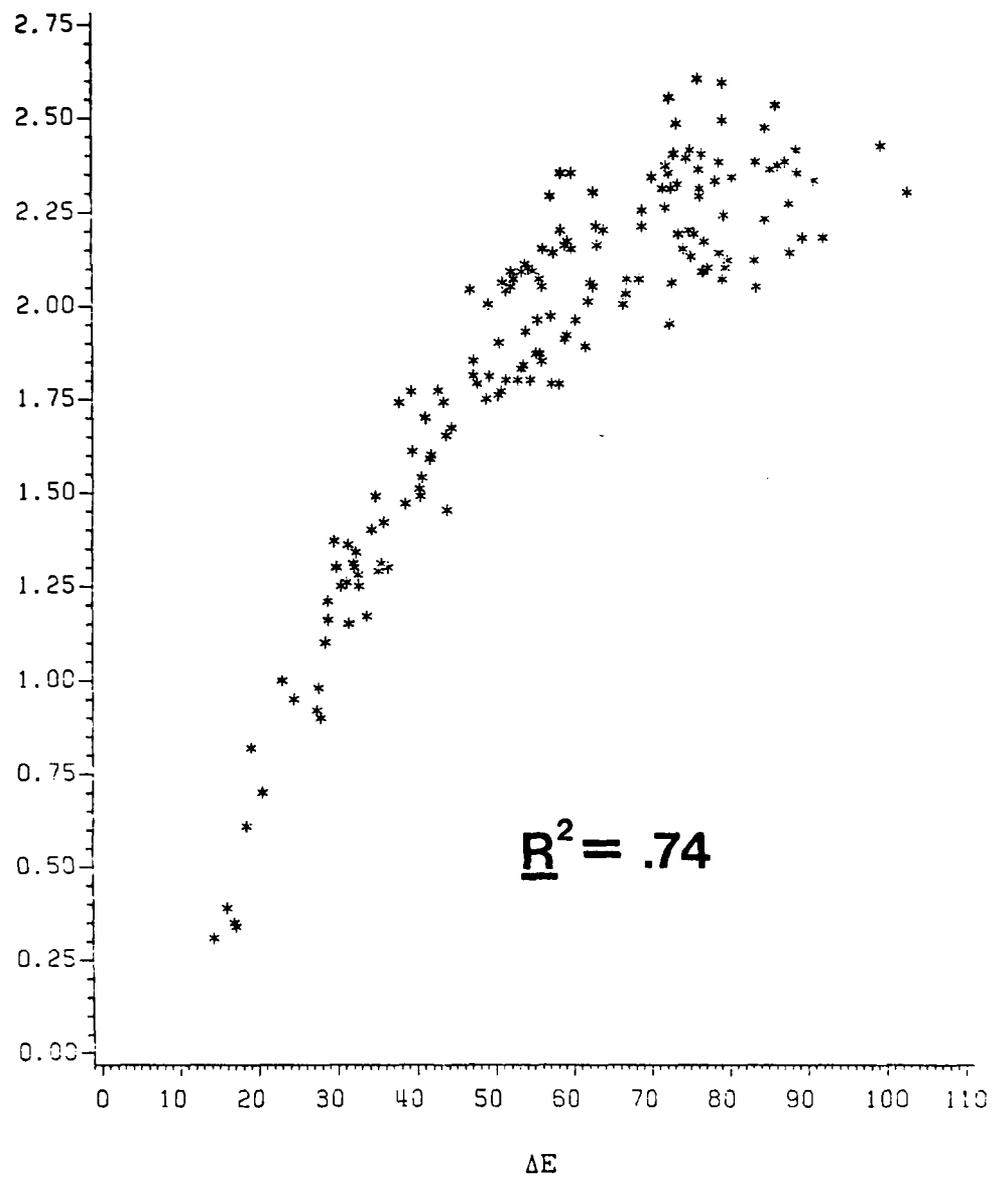


Figure 22.  $\Delta E$  in luminance-generalized,  $\Delta E$ -standardized ( $L^*, u^*, v^*$ ) space by RS scattergram, Study 2 (Lippert, 1985).

equivalently, a displayed chrominance-difference of 0.6 (on the 2.2:1 rescaled  $u':v'$  diagram). The 100% legibility predictions are for contrast conditions greater than any tested in either study because asymptotic performance was obtained with combinations of the highest luminance ratios and chrominance-differences tested.

The generalized, standardized metric expressions are listed in Table 3, which includes each metric's color component weightings as described in the INTRODUCTION. It is important to note that the  $\Delta E$  metrics describe RS consistently across the 0.5 m and 0.76 m viewing distances, although they could not be expected to do so for any viewing distance, character size, dot width, etc. It might be found through further experimentation that the luminance contrast to chrominance contrast weighting, R, while determined to be constant for Studies 1 and 2, might vary with the spatial content of the imagery of interest.

A somewhat simpler form of metric will now be considered. Researchers (e. g., Boynton, 1981, and others) working on the visual distinctness of bordering multicolor fields and the physiology of primate vision have suggested that the difference in stimulation rate of the long-wavelength receptors at a T/B border,  $\Delta X$ , might better describe RS with isoluminous T/B pairs than does  $\Delta E$ . RS for the averaged  $M = 0.006$  conditions in Study 1 were subjected to regression

Table 3.  $\Delta E$  Component Weightings and  
Luminance-generalized,  $\Delta E$ -standardized  
Expressions of the Study 1 Metrics:

- (a).  $Y, u', v'$
- (b).  $\log Y, u', v'$
- (c).  $L^*, u', v'$
- (d).  $L^*, u^*, v^*$

$\Delta E$  Component Weightings:

- (a).  $((\Delta Y)^2 + (110\Delta u')^2 + (50\Delta v')^2)^{0.5}$
- (b).  $((\log(Y_T/Y_B))^2 + (1.43\Delta u')^2 + (0.65\Delta v')^2)^{0.5}$
- (c).  $((\Delta L^*)^2 + (79\Delta u')^2 + (36\Delta v')^2)^{0.5}$
- (d).  $((\Delta L^*)^2 + (0.0058\Delta u^*)^2 + (0.0028\Delta v^*)^2)^{0.5}$

Generalized, Standardized Expressions:

- (a).  $((155/Y_M)\Delta Y)^2 + (367\Delta u')^2 + (167\Delta v')^2)^{0.5}$
- (b).  $((250/\log Y_M)(\log(Y_T/Y_B)))^2 + (7.67\Delta u')^2 + (3.49\Delta v')^2)^{0.5}$
- (c).  $((260/L_M^*)\Delta L^*)^2 + (440\Delta u')^2 + (200\Delta v')^2)^{0.5}$
- (d).  $((595/L_M^*)\Delta L^*)^2 + (0.75\Delta u^*)^2 + (0.36\Delta v^*)^2)^{0.5}$

analysis with  $\Delta X$ , resulting in  $r^2 = 0.52$ , compared to  $r^2 = 0.74$  for a 2.2:1 ( $u':v'$ ) rescaled metric. For Study 2 isoluminous pairs,  $r^2 = 0.54$  and  $0.75$  for the  $\Delta X$  and rescaled ( $u',v'$ ) metrics, respectively.

There are slight amounts of luminance contrast among the most nearly isoluminous pairs tested and  $\Delta X$  does incorporate some luminance contrast information, while the ( $u',v'$ ) metric does not.  $\Delta X$  is certainly a more parsimonious means of RS prediction, and it works reasonably well for the isoluminous case, but not as well as rescaled ( $u',v'$ ). For rescaled ( $Y,u',v'$ ),  $r^2 = 0.74$  and  $0.75$ , respectively, for Studies 1 and 2, differing from the  $r^2$ s for rescaled ( $u',v'$ ) only beyond the second decimal place. When used to describe the entire Study 1 data set,  $r^2 = 0.18$  for  $\Delta X$ , indicating the limit of its effectiveness compared to rescaled  $\Delta E(Y,u',v')$ , for which the simple linear  $r^2 = 0.49$ . For the complete Study 2 data set,  $r^2 = 0.20$  versus  $r^2 = 0.54$  for  $\Delta X$  and rescaled  $Y,u',v'$  space, respectively.

' For color contrast including luminance contrast, a two-component formulation might be considered in which the root sum of squares of  $\Delta Y$  and  $\Delta X$  are represented. For RS in Study 1, a 2.2:1 Y:X rescaling is required for the maximum obtainable  $r^2 = 0.43$  and, for Study 2,  $r^2 = 0.45$ , which compare reasonably well with the linear  $Y,u',v'$  models in the preceding paragraph.

While it appears there is some loss of descriptive information in the rescaled (Y,X) metric relative to the rescaled (Y,u',v'), (Y,X) is a reasonably good predictor of RS. A rescaled (Y,X) metric, comprised of  $\Delta E/\Delta E^2/N$  terms, as those reported in Table 2, results in  $R^2 = 0.84$  as compared to the  $R^2 = 0.92$  for rescaled (Y,u',v') with respect to the entire Study 1 data set. For Study 2,  $R^2_s = 0.85$  and  $0.95$  for the three-term rescaled (Y,X) and (Y,u',v') metrics, respectively.

The rescaled (u',v') combinations with Y, logY, or L\* are each superior to either the (X) or (Y,X) metrics in the accuracy of their description of RS whether luminance contrast is present or not. The single rescaling ratio of 2.2:1 for u':v' and Y:X to obtain maximum  $r^2_s$  is ruled a coincidence unless further study shows otherwise.

In summary, the results of the two studies are consistent and conclusive. Several, simple, first-order  $\Delta E$  scales exist which serve equally well to describe or predict RS with multicolor CRT raster imagery for a range of character luminances, over two viewing distances, and in both positive and negative presentation polarities. These are the Y,u',v', logY,u',v', L\*,u',v', and L\*,u\*,v\* rescaled color spaces (Table 3). Because of its predictive accuracy and simplicity, a luminance-generalized,  $\Delta E$ -standardized Y,u',v' metric, accounting for 71% and 75% of the RS variability in

Studies 1 and 2, respectively, is recommended as the most appropriate metric of emissive display legibility to be tested in these studies.

Even greater predictive power can be obtained with a second-order equation using the rescaled axes, as illustrated in Table 2. Again, the  $Y,u',v'$  space is slightly superior in its predictive power and is recommended due to its non-convergent geometry and its inherent simplicity. The second-order equations predict, on the average, about 20 percent more of the RS than do the first-order equations. Thus, while the first-order equations are simpler and more conservative, the second-order equations are recommended, within the limits of the present research parameters, for more precise prediction.

## APPLICATIONS

The lack of statistical dependency between the  $\Delta E$  and N effects suggests that for many information extraction tasks, the  $\Delta E$  metrics described might perform operationally in an accurate relative manner. That is, factors other than color contrast might be found to contribute in constant, additive fashions to task performance speed, as did N. Unfortunately, logic dictates that such realistic attributes as image dynamics and character surround spatial complexity must, by definition, alter the simple specifications of color contrast permitted by this laboratory study and, therefore, cannot be independent of  $\Delta E$ . To date, the only known experiment to test the legibility of numerals displayed against spatially complex multicolor backgrounds with a shadowmask CRT was conducted by Post (1983).  $\Delta E$  scales (from the convergent and non-convergent spaces tested) based on Post's indirectly determined color parameters, do not describe Post's measure of RS with static imagery. It appears that dynamic imagery would be an even less likely candidate for legibility modeling  $\Delta E$  metrics of the types considered in this report.

However, it is hoped that the legibility metrics reported here will find operational confirmation in daylight color CRT environments, airborne head-down displays, etc. In such cases, displayed color parameters will have to be corrected

for the ambient illumination, which may vary spatio-temporally and require the estimation of "worst case" conditions.

In a cockpit environment, for example, RS might be determined for a relatively small sample of color contrast conditions spanning at least the luminance contrast range tested in the present research. If  $\Delta E$  versus RS regression analyses prove similar to those reported here, it would be reasonable to assume a global, or uniform, property for the metric in the new application.

There are many task environments such as CAD/CAM stations and office VDTs where the legibility metrics reported here should find direct application. Because numeral reading accuracy was high across the range of numeral reading speed, it might be possible to employ increasing steps in  $\Delta E$  to color displayed information one wishes to separate into increasing levels of importance or salience.

Although the present findings relate specifically to the legibility of CRT raster imagery, it is suggested that the applicability of the  $\Delta E$  metrics be tested for discrete picture element electronic imaging technologies such as liquid crystal displays (LCDs) and full-color electroluminescent displays. In electronic display viewing environments in which the individual picture elements are

not readily distinguishable by the user, that is, in well designed viewing conditions, it might prove that spatial content differences between discrete element and CRT raster images do not significantly affect legibility. In such a case, the metrics reported here should apply to the discrete element display technology, at least across a 10-50 cd/m<sup>2</sup> image luminance range.

Where  $\Delta E$  is shown to act independently in its effect on visual performance, experimentation might lead to operational metrics which could be employed in reverse. Time spent on a visual information extraction task might be "metered" through the specification of color contrast to obtain the desired (i. e., reduced) level of performance, effectively loading operator faculties a known amount.

## CONCLUSIONS

There are several important conclusions to be drawn from this work. In visual environments where electronically displayed contrast between symbology and uniform backgrounds is limited to a luminance modulation less than 0.27 (i. e., a luminance contrast ratio less than 1.6), legibility can be enhanced by employing either characters or backgrounds of red and magenta hues. This finding was obtained in Study 1, where relatively high levels of RS were found for red numerals, no matter the background hue, even without luminance contrast. Overall, a luminance contrast ratio of 1.6 might be considered minimally adequate under ideal viewing conditions, those free of vibration, glare, luminance adaptation effects, etc.

It is interesting that the suprathreshold performance investigated in the present work is well modeled by a chrominance space altered only slightly from one determined to be the best available for threshold modeling. In many areas of science, a disagreement between empirical findings in matching theoretical models of 2:1 is considered trivial. However, the  $(u',v')$  construct is an empirical development and the 2.2:1 disparity found in this research questions the adequacy of existing colorimetric methodologies and argues for continued research toward a more coherent color science of visual perception and performance.

The CIE 1976  $L^*,u^*,v^*$  metric is presently finding general application as a predictor of many types of color visual behavior, largely because it has been the available standard since its inception and is empirically validated to be reasonably uniform with respect to threshold levels of color contrast containing little or no luminance contrast. The present research has determined that convergent metrics such as  $L^*,u^*,v^*$  are inappropriate bases for the prediction of any visual performance measure which is monotonically related to luminance contrast. However, it has also been demonstrated that the  $L^*,u^*,v^*$  axes, if linearly, independently rescaled, yield a metric of RS predictive power similar to that of the non-convergent metrics reported. In fact, the great expansion of  $L^*$  in rescaled  $L^*,u^*,v^*$  indicates that the statistical rescaling procedure increases the effect of luminance contrast in the rescaled  $\Delta E$  metric and serves to reduce the rescaled space's rate of convergence with respect either to luminance or chrominance units.

Color difference in non-convergent 3-spaces always varies monotonically with luminance difference. Appropriate non-convergent combinations of any of several luminance scalings with rescaled  $(u',v')$  provides  $\Delta E$  metrics which describe RS well for both positive and negative presentation polarities, simplifying the image engineering process.

The luminance-generalized,  $\Delta E$ -standardized expressions of all of the metrics developed through this work yield nearly identical predictions in terms of percent improvement in the legibility of CRT raster imagery and perhaps will be shown to apply to other emissive display imagery. The  $\Delta E$  formulations in Table 3 are designed specifically for emissive displays. The non-convergent spaces reported here do not include the specification of illuminants, etc., which are meaningless in pure self-luminous image characterization and require additional computations to yield measures of  $\Delta E$ . Because of its simplicity and RS predictive power,  $Y, u', v'$  space in a luminance-generalized,  $\Delta E$ -standardized form is the most highly recommended design aid for legible multicolor CRT raster imagery to come from these studies.

A 2.2:1 ( $u', v'$ ) rescaling results in chrominance differences which uniformly describe RS with all of the isoluminous color contrast conditions tested. The ( $u', v'$ ) rescaling ratio remains 2.2:1, in combination with any of the luminance scalings tested, to describe RS for the color contrast conditions including luminance contrast. This is one argument for a model of color visual recognition in which luminance contrast and chrominance contrast components, as defined by the CIE, are independent in their effects. Given the energies previously spent investigating various luminance x chrominance interaction effects inherent

AD-A175 062

UNITARY SUPRATHRESHOLD COLOR-DIFFERENCE METRICES OF  
LEGIBILITY FOR CRT RA. (U) VIRGINIA POLYTECHNIC INST  
AND STATE UNIV BLACKSBURG HUMAN FAC.  
T M LIPPERT ET AL. SEP 86 HFL/ONR-86-3

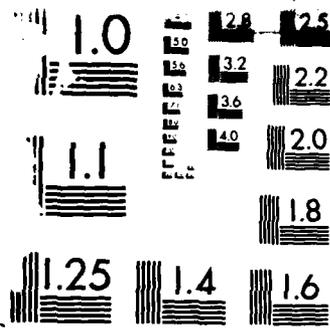
2/2

UNCLASSIFIED

F/G 5/10

NL





MICROCOPY RESOLUTION TEST CHART  
NBS 1963-A

to color- and brightness-matching performance, which were briefly discussed in the INTRODUCTION, the apparent lack of similar effects on RS deserves further study.

Perhaps most importantly, the assumption that the contribution to legibility of a specific combination of luminance contrast and chrominance-difference units is invariant with respect to the absolute level of retinal stimulation is supported for stimulus combinations ranging in luminance from 10 to 50  $\text{cd/m}^2$ . This contrast "constancy" effect might be shown to generalize to higher luminance imagery through continued effort.

## REFERENCES

- Booker, R. L. Luminance-brightness comparisons of separated circular stimuli. Journal of the Optical Society of America. 1981, 71, 139-144.
- Carter, E. C. and Carter, R. C. Color and conspicuousness. Journal of the Optical Society of America, 1981, 71, 723-729.
- Coblentz, W. W. and Emerson, W. B. Relative sensibility of the average eye to light of different colors and some practical applications of radiation problems. U.S. Bureau of Standards Bulletin, 1917, 14, 167-236.
- Cohen, J. and Friden, T. P. The Euclidean nature of color space. Bulletin of the Psychonomic Society, 1975, 5, 159-161.
- Costanza, E. B. An evaluation of a method to determine suprathreshold color contrast on CRT displays. Unpublished Master's Thesis, Virginia Polytechnic Institute and State University, May, 1981.
- Farley, W. W. and Gutmann, J. C. Digital image processing systems and an approach to the display of colors of specified chrominance. Virginia Polytechnic Institute and State University Technical Report

HFL-80-2/ONR-80-2, August, 1980.

Frome, F. S., Buck, S. L., and Boynton, R. M. Visibility of borders: separate and combined effects of color differences, luminance contrast, and luminance level, Journal of the Optical Society of America, 1981, 71, 2, 145-150.

Galves, J-P. and Brun, J. Color and brightness requirements for cockpit displays: proposal to evaluate their characteristics. Twenty-ninth AGARD Avionics Panel Technical Meeting, 1981.

Gibson, K. S. and Tyndall, E. P. T. Visibility of radiant energy. U.S. Bureau of Standards, Scientific Paper No. 475, 1923, 19, 131-191.

Graham, C. H. Color: data and theories. In C. H. Graham (Ed.) Vision and visual perception. New York: Wiley, 1965.

Grassman, H. On the theory of compound colors. Philosophy Magazine, 1854, 7, 254-264.

Guild, J. The colorimetric properties of the spectrum. Philosophical Transactions of the Royal Society, 1931, 230A, 149-187.

Hunt, R. W. G. The specification of colour appearance.

I. Concepts and terms. Color, 1977, 2, 55-59.

Lippert, T. M. Color contrast effects for a simulated CRT HeadUp Display. Unpublished Master's Thesis, Virginia Polytechnic Institute and State University, 1984.

MacAdam, D. L. Visual sensitivities to color differences in daylight. Journal of the Optical Society of America, 1942, 32, 247-274.

Post, D. L., Costanza, E. B., and Lippert, T. M. Expressions of color contrast as equivalent achromatic contrast. Proceedings of the 26th Annual Human Factors Society Meeting, 1982, 581-585.

Post, D. L., Lippert, T. M., and Snyder, H. L. Quantifying color contrast. Proceedings of the Society of Photo Optical Engineers. 1983, 12-19.

Post, D. L. Color contrast metrics for complex images. Unpublished Doctoral Dissertation, Virginia Polytechnic Institute and State University, 1983.

Snyder, H. L. Human visual performance and flat panel display image quality. Virginia Polytechnic Institute and State University Technical Report HFL-80-1/ONR-80-1, July, 1980.

Statistical Analysis System User's Guide: Statistics.

Alice Allen Ray (Ed) 1982, 139-199.

Wyszecki, G. and Stiles, W. S. Color science. New York:

Wiley, 1967, 1982.

APPENDIX A. Stimulus Color Parameters:  
 Studies 1 and 2, (Lippert, 1984, 1985)

Table 4. Stimulus Color Parameters: Study 1.

---1931 CIE---					
Color Name	Code	M Level	x	y	Y (cd/m <sup>2</sup> )
Numerals (T)					
Achromatic	ACH	all	.310	.316	46.3
Yellow-green	Y-G	all	.370	.472	45.5
Red	RED	all	.605	.344	48.1
Backgrounds (B)					
Purple	p	.006	.383	.222	47.9
		.060	.378	.220	42.8
		.120	.380	.220	38.6
		.164	.380	.219	34.9
		.210	.372	.220	31.0
		.270	.373	.219	27.5
Violet	v	.316	.374	.218	25.0
		.006	.253	.202	46.0
		.060	.255	.191	42.6
		.120	.254	.190	38.4
		.164	.255	.190	34.5

Table 4, continued.

		.210	.253	.189	31.1
		.270	.255	.189	27.8
		.316	.255	.190	25.0
Blue	b	.006	.243	.264	45.3
		.060	.243	.264	40.8
		.120	.242	.263	36.6
		.164	.242	.262	32.8
		.210	.243	.265	29.6
		.270	.241	.261	26.3
		.316	.241	.262	23.6
Green	g	.006	.267	.438	44.1
		.060	.265	.435	39.1
		.120	.265	.434	34.5
		.164	.264	.434	31.3
		.210	.264	.434	28.2
		.270	.264	.433	25.4
		.316	.266	.436	22.9
Yellow	y	.006	.412	.404	45.7
		.060	.423	.421	40.5
		.120	.423	.419	36.5
		.164	.422	.417	33.0
		.210	.425	.419	29.7
		.270	.408	.406	26.4

Table 4, continued.

		.316	.423	.419	23.9
Orange	o	.006	.389	.367	45.8
		.060	.387	.366	41.0
		.120	.382	.367	36.8
		.164	.387	.365	33.5
		.210	.386	.364	30.0
		.270	.385	.364	27.0
		.316	.386	.362	24.1
Red	r	.006	.440	.344	47.7
		.060	.440	.344	42.6
		.120	.433	.340	38.5
		.164	.434	.344	34.6
		.210	.437	.341	30.4
		.270	.436	.342	27.4
		.316	.429	.340	24.9
Achromatic	a	.006	.308	.314	45.7
		.060	.308	.311	41.0
		.120	.306	.310	36.8
		.164	.309	.310	33.4
		.210	.304	.319	33.2
		.270	.308	.310	26.9
		.316	.305	.308	24.0

Table 5. Stimulus Color Parameters: Study 2.

---1976 CIE---					
Color Name	Code	M Level	u'	v'	Y (cd/m <sup>2</sup> )
Numerals (T)					
Achromatic	ACH	all	.200	.480	20.2
Yellow-green	Y-G	all	.180	.540	20.1
Red	RED	all	.350	.480	20.0
Backgrounds (B)					
Achromatic	a	.290	.185	.481	11.2
		.181	.196	.475	14.0
		.083	.202	.470	16.8
		.005	.207	.477	20.1
		.089	.211	.483	23.7
		.185	.191	.486	29.5
Green	g	.270	.209	.470	35.2
		.290	.141	.502	11.1
		.181	.143	.495	13.8
		.083	.145	.510	17.0
		.005	.135	.516	20.2
		.089	.133	.488	23.9

Table 5, continued.

		.185	.146	.510	28.7
		.270	.153	.508	35.7
Red	$r_1$	.290	.273	.469	11.3
		.181	.281	.477	13.8
		.083	.270	.482	17.1
		.005	.255	.465	20.0
		.089	.264	.445	24.3
		.185	.283	.430	28.8
		.270	.279	.458	33.9
Red	$r_2$	.290	.305	.491	10.8
		.181	.301	.482	13.9
		.083	.293	.485	16.9
		.005	.310	.474	18.8
		.089	.305	.480	23.7
		.185	.294	.466	30.1
		.270	.287	.476	34.7
Magenta	$m_1$	.290	.236	.435	11.1
		.181	.251	.444	13.8
		.083	.233	.428	16.8
		.005	.249	.437	20.0
		.089	.242	.431	24.1
		.185	.253	.449	29.4

Table 5, continued.

		.270	.230	.422	34.2
Magenta	$m_2$	.290	.273	.411	11.1
		.181	.286	.395	14.0
		.083	.301	.404	17.2
		.005	.305	.402	20.0
		.089	.294	.418	24.0
		.185	.283	.400	29.2
		.270	.911	.413	34.6
Purple	$p_1$	.290	.211	.393	11.0
		.181	.207	.405	13.7
		.083	.196	.411	17.1
		.005	.191	.384	19.9
		.089	.220	.399	24.4
		.185	.208	.390	29.4
		.270	.212	.408	35.6
Purple	$p_2$	.290	.245	.303	10.9
		.181	.217	.288	13.8
		.083	.237	.310	17.2
		.005	.241	.319	20.3
		.089	.233	.315	24.2
		.185	.209	.323	28.6
		.270	.221	.301	35.9

END

1-87

Dttic